

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
“КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМ. ІГОРЯ СІКОРСЬКОГО ”  
Факультет електроніки  
Кафедра електронної інженерії

До захисту допущено  
Завідувач кафедри  
В. І. Тимофєєв  
“ ” 20\_\_ р.

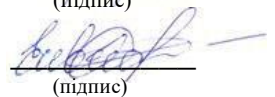
**Дипломна робота**  
освітнього рівня «бакалавр»  
за спеціальністю 153 мікро- та наносистемна техніка

на тему «Нанопорове секвенування ДНК для виявлення генетичних змін»

Виконала студентка 4 курсу, групи ДМ-61  
Євдощенко Ірина Миколаївна  
(прізвище, ім'я, по батькові)

  
(підпис)

Керівник доц. каф. ЕІ, доц., к. т. н. Іванько К. О.  
(посада, вчене звання, науковий ступінь, прізвище та ініціали)

  
(підпис)

Рецензент доц. каф. АМЕС, доц., к. т. н. Попович П. В.  
(посада, вчене звання, науковий ступінь, прізвище та ініціали)

  
(підпис)

Засвідчую, що у цій дипломній роботі немає  
запозичень з праць інших авторів без  
відповідних посилань.

Студент   
(підпис)

**Національний технічний університет України**  
**“Київський політехнічний інститут імені Ігоря Сікорського”**

Факультет електроніки  
 Кафедра електронної інженерії  
 Освітній рівень «бакалавр»  
 за спеціальністю 153 мікро- та наносистемна техніка

**ЗАТВЕРДЖУЮ**

**Завідувач кафедри**

\_\_\_\_\_ В. І. Тимофєєв  
 “ \_\_\_\_ ” \_\_\_\_\_ 20\_\_ р.

## ЗАВДАННЯ

### НА ДИПЛОМНУ РОБОТУ

Євдощенко Ірині Миколаївні

(прізвище, ім'я, по батькові)

1. Тема роботи «Нанопорове секвенування для визначення генетичних змін»

керівник роботи Іванько Катерина Олегівна, к. т. н, доц. каф. ЕІ

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету від “ \_\_\_\_ ” \_\_\_\_\_ 20\_\_ року № \_\_\_\_\_

2. Строк подання студентом роботи: 12 червня 2020 року.

3. Вихідні дані до роботи: моделювання сигналів по нуклеотидній послідовності ДНК, знешумлення сигналів, візуалізації дослідження, класифікація сигналів

4. Зміст дипломної роботи (перелік питань, які потрібно розробити): вивчення методів нанопорового секвенування, дослідження змін в ДНК, класифікація сигналів нанопорового секвенування, дослідження методів вимірювання потенціалу на мембрані, створення програмного забезпечення для: моделювання сигналів по заданій послідовності, пошуку подібних областей послідовності методом кореляційного аналізу, очистки сигналів від шумів, класифікації сигналів, візуалізація досліджень.

5. Перелік графічного (ілюстративного) матеріалу (із зазначенням обов'язкових креслень, плакатів, презентацій тощо): слайди презентації для захисту.

6. Дата видачі завдання: «13» квітня 2020 року

## КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів виконання дипломної роботи	Строк виконання етапів роботи	Примітка
1	Підбір літератури з теми	13 квітня	
2	Пошук баз даних, вивчення особливостей патологій, властивих раку	20 квітня	
3	Вивчення методу нанопорового секвенування для визначення послідовностей ДНК, аналіз методів вимірювання потенціалів на мембрані	27 квітня	
4	Розробка і тестування програмних засобів, висновки по роботі	4 травня	
5	Оформлення пояснювальної записки	31 травня	
6	Проходження перевірки на плагіат та нормоконтроль	1 червня	

Студент

  
 (підпис)
Євдощенко І. М.  
(прізвище та ініціали)

Керівник роботи

  
 (підпис)
Іванько К. О.  
(прізвище та ініціали)

## РЕФЕРАТ

Пояснювальна записка: 82 с., 4 ч., 36 рис, 42 джерела

ДНК, НАНОПОРОВЕ СЕКВЕНУВАННЯ, МЕТОДИ МОДЕЛЮВАННЯ  
ТА АНАЛІЗУ СИГНАЛІВ.

Об'єктом розгляду є послідовність нуклеотидів молекули ДНК. Предметом дослідження є застосування методів моделювання та аналізу сигналів отриманих шляхом нанопорового секвенування.

Метою роботи є виявлення генетичних змін в послідовності, очищення сигналів від шумів, класифікація сигналів нанопорового секвенування.

В першому розділі наведено огляд літератури стосовно ДНК і мутацій: чому вони стаються, як їх виявляють і класифікують. В другому розділі інформація про нанопорове секвенування: процес секвенування, типи пор, які застосовуються, варіанти нанпорового секвенування та протоколи. В третьому розділі наведено методи вимірювання потенціалу на мембрані, наприклад, метод петч –кламп, де фіксується напруга і реєструється зміна іонного струму. Наведено приклад аналізу сигналів, а саме кореляційний аналіз, з метою знаходження однакових фрагментів послідовності ДНК. В четвертому розділі представлено моделювання сигналів різних класів (послідовність в нормі і з мутацією), методи фільтрації для усунення шумів сигналу та класифікація послідовностей методом К – найближчих сусідів.

## ABSTRACT

Explanatory note 82 p., 4 s., 36 figures, 42 references.

DNA, NANOPORE SEQUENCE METHODS OF MODELING AND ANALYSIS OF SIGNALS.

The object of consideration is the nucleotide sequence of the DNA molecule.

The subject of the research is the application of methods of modeling and analysis of signals obtained by nanoporous sequencing.

The purpose of the work is to identify genetic changes in the sequence, purification of signals from noise, their classification.

The first section reviews the literature on DNA and mutations: why they occur, how they are detected and classified. The second section provides information on nanoporous sequencing: the sequencing process, the types of pores used, nanoporous sequencing options, and protocols. The third section presents methods for measuring the potential on the membrane, for example, for our work it is the method of patch-clamps, where the voltage is fixed and the change in ionic current is registered. An example of signal analysis, namely correlation analysis, in order to find the same fragments of the DNA sequence. The fourth section presents the modeling of signals of different classes (normal and mutated sequence), filtering methods to eliminate signal noise and classification of sequences by the method of K - nearest neighbors.

## ЗМІСТ

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ .....	8
ВСТУП.....	9
1 МУТАЦІЇ. КЛАСИФІКАЦІЯ МУТАЦІЙ. ПРОЦЕС ПЕРЕТВОРЕННЯ НОРМАЛЬНОЇ КЛІТИНИ В РАКОВУ.....	10
1.1 Структура ДНК.....	10
1.2 Ракові клітини. Мутації, що призводять до раку .....	12
1.3 Порівняльне секвенування.....	19
1.4 Помірні і рідкі онкогени .....	20
1.5 Гени достворіно мутуючі при раку .....	21
1.6 Рак стравоходу.....	24
1.7 Рак молочної залози .....	26
1.8 Висновок до розділу 1 .....	29
2 МЕТОД НОВОГО ПОКОЛІННЯ – НАНОПОРОВЕ СЕКВЕНУВАННЯ ДНК .....	30
2.1 Біологічні нанопори .....	31
2.2 Як працюють нанопори при секвенуванні ДНК.....	31
2.4 Екзонуклеазне секвенування .....	36
2.5 Типи нанопор.....	36
2.5.1 Білкові нанопори .....	36
2.5.2 Твердотільні нанопори .....	40
2.5.3 Порівняння біологічних і твердотільних нанопор.....	41
2.6. Основні протоколи секвенування ДНК.....	42
2.6.1. Протокол Rapid .....	42
2.6.2. Протокол 1D .....	43
2.6.3 Протокол 1D <sup>2</sup> .....	45
2.7. Виділення довгої ДНК .....	46
2.8. Яким чином нанопорове секвенування допоможе при вивченні раку.....	47

2.9 Висновки до розділу 2 .....	48
3 МЕТОДИ ВИМІРЮВАННЯ МЕМБРАННОГО ПОТЕНЦІАЛУ .....	49
3.1 Метод фіксації потенціалу .....	49
3.2 Петч – кламп метод .....	50
3.3 Висновки до розділу 3 .....	53
4 РОЗРОБКА АЛГОРИТМІВ ДЛЯ МОДЕЛЮВАННЯ ТА АНАЛІЗУ СИГНАЛІВ НАНОПОРОВОГО СЕКВЕНУВАННЯ ДНК .....	54
4.1 Пошук подібних областей між двома і більше послідовностями методом кореляційного аналізу.....	63
4.2 Алгоритм знешумлення сигналів. ....	68
4.3 Класифікація сигналів нанопорового секвенування .....	71
4.4 Висновки до розділу 4 .....	75
ВИСНОВКИ .....	77
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ .....	78

## СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ

АТФ – аденозинтрифосфат

ДНК – дезоксирибонуклеїнова кислота

ПАР – поверхневі активні речовини

ПЛР – полімеразна ланцюгова реакція

РНК – рибонуклеїнова кислота

COSMIC – catalogue of somatic mutation in cancer (раковий каталог соматичних мутацій)

DCIS – ductal carcinoma in situ (карцинома проток, в раку молочної залози)

IDC – invasive ductal carcinoma (інвазивна карцинома проток, в раку молочної залози)

NCBI – national Center for biology information (національний центр біотехнологічної інформації)

NGS – next generation sequencing (секвенування нового покоління)

OSCC – oesophageal squamous cell carcinoma (плоскоклітинна карцинома стравоходу)

pH – водневий показник, вказує на ступінь кислотності або лужності розчину

SNV –Single Nucleotide Variation (варіація одиничного нуклеотиду)

TCGA – cancer Genome Atlas Program (атлас ракового геному)

TSCS – Topographic Single Cell Sequencing (топографічне секвенування окремих клітин)



## ВСТУП

Нанопорове секвенування на сьогодні відомо як провідний інструмент генної інженерії. Завдяки прочитанню послідовностей ДНК ми можемо виявити патогенну зміну ще до того як це перетече в хворобу. Раннє діагностування вкрай важливе. На сьогоднішній день вже є технології редагування геному і в недалекому майбутньому вони стануть доступними для кожного. Для того, щоб відредагувати мутований фрагмент необхідно спочатку знайти його розташування. Дана робота націлена на виявлення мутацій, а саме мутацій характерних раку, в певній ділянці ДНК. Україна є другою країною в світі по кількості захворювання на рак, за останні 10 років населення скорилось на 4 мільйони через онкологічні захворювання. З цією проблемою можна і потрібно боротися і вчасне діагностування чи то вже запущеного організмом поділу онкологічних клітин, чи виявлення високої ймовірності виникнення раку може врятувати мільйони людей.

Метою роботи є вивчення процесу нанопорового секвенування, дослідження змін в ДНК, що призводять до раку, методів, якими користуються для вимірювання струму при проходженні ДНК через пору, розробка алгоритмів для моделювання і аналізу сигналів, класифікація сигналів.

У роботі будуть представлені: література, методи, що застосовуються для аналізу поставленої задачі, моделювання та аналіз сигналів нанопорового секвенування, демонстрація отриманих результатів, висновки по роботі.

# 1 МУТАЦІЇ. КЛАСИФІКАЦІЯ МУТАЦІЙ. ПРОЦЕС ПЕРЕТВОРЕННЯ НОРМАЛЬНОЇ КЛІТИНИ В РАКОВУ

## 1.1 Структура ДНК

За своєю хімічною будовою ДНК є біополімером, мономерами якого являються нуклеотиди. Молекула ДНК зазвичай складається з двох ланцюгів спірально закручених і з'єднаних між собою воднем, що утворюється між азотистими основами. Кожний нуклеотид складається з: азотистої основи, дезоксирибози, залишку фосфорної кислоти (рис. 1.1). В свою чергу в ДНК входять наступні азотисті основи: аденін, гуанін, тимін і цитозин. Азотисті основи одного ланцюга молекули ДНК з'єднується з азотистою основою другого суворо по принципу компліментарності: аденін тільки з тиміном, гуанін тільки з цитозином [1]. ДНК негативно заряджена завдяки заряду фосфатних груп. Від'ємний заряд фосфатних груп зберігає молекули всередині, саме він не дозволяє фосфатним ефірам піддаватися процесу гідролізу та зберігає стійкість нуклеїнових кислот [2].

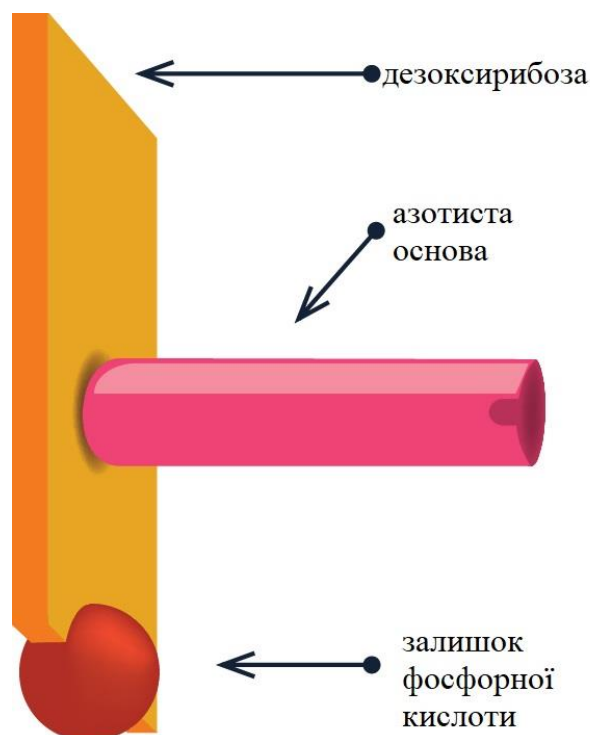
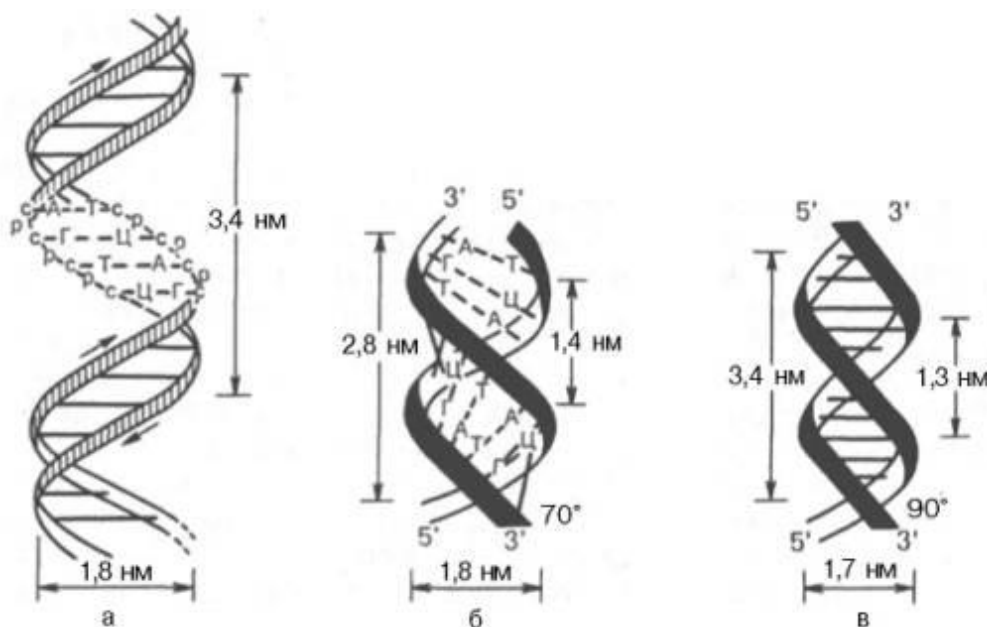


Рисунок 1.1 – Будова нуклеотиду

Обидва ланцюга полі – нуклеотидів в біспіральній молекулі ДНК мають суворо визначене просторове розташування, при якому азотисті основи знаходяться всередині, а фосфорильні і вуглеводні компоненти – зовні [3].

Крок спіралі, тобто відстань між витками, складає 3,4 нм (рис.1.2 а). На цій ділянці розміщується 10 нуклеотидних залишків, розмір нуклеотиду складає 0,34 нм, діаметр біспіральної молекули 1,8 нм. Конфігурація подвійної спіралі ДНК сильно змінюється в залежності від кількісного складу води та іонної сили оточуючого середовища.

На (рис.1.2 б, в) наведено дві прості форми (А – форма і В – форма) конфігурації ДНК відповідно [3].



с – залишок дезоксирибози, р – залишок фосфорної кислоти.

Рисунок 1.2 – Схематичне зображення подвійної спіралі [3]

Обидва ланцюга в молекулі ДНК мають протилежну полярність. Це означає, що міжнуклеотидний зв'язок в одному ланцюзі має напрямок  $5' \rightarrow 3'$ , а в інший  $3' \rightarrow 5'$  [3]. Початком ланцюга вважають  $5'$  - кінець, у якого  $5'$  - вуглецевий атом цукру має тільки фосфорний залишок. Закінчення ланцюга – це нуклеотид, у якого  $3'$  - вуглецевий атом цукру містить  $\text{OH}$  – групу. Цей нуклеотид називають  $3'$  - кінцем полінуклеотидного ланцюга. Розрізняти початок і кінець полінуклеотидного ланцюга принципово важливо,

так як при описі послідовності нуклеотидних залишків генетичного коду їх перераховують від 5' - кінця до 3' - кінця [4].

Генетичний кодом називають систему запису генетичної інформації про послідовність розміщення амінокислот в білках у вигляді послідовності нуклеотидів РНК або ДНК. Кожній амінокислоті білка відповідає послідовність з трьох розташованих один за одним нуклеотиди ДНК – триплет (кодон). Кожний триплет нуклеотидів кодує певну амінокислоту, яка буде вбудована в поліпептидний ланцюг [5].

Амінокислоти поділяються на замінні і незамінні, де незамінними амінокислотами називають ті амінокислоти, які організм не синтезує. Таких амінокислот є дев'ять з двадцяти, вони потрапляють в організм з білковою їжею. Рівень певних амінокислот в організмі впливає на поділ ракових клітин.

Наприклад, лейцин – амінокислота, що входить до складу великої кількості білків, грає важливу роль в розвитку стійкості до тамоксифену (блокує рівень естрогену в ракових клітинах) при естрогеннопозитивному раку молочної залози. Зниження рівня лейцину гальмує поділ пухлинних клітин, а збільшення, навпаки, сприяло клітинному поділу [6]. Амінокислота глютамін – ракові клітини при поділі використовують глютамін як «топливо» для свого росту.

## 1.2 Ракові клітини. Мутації, що призводять до раку

Мутація - будь-яка зміна в молекулі ДНК або структурі хромосом. Зміни, які відбуваються в гаметах - статевих клітинах, називають гермінальними, в будь-яких інших клітинах організму - соматичними. Гермінальні мутації передаються від батьків у спадок у вигляді генетичних захворювань. Соматичні зміни не успадковуються, а виникають в результаті зовнішніх впливів (куріння, радіоактивного випромінювання та інших). Доведено, що злоякісні пухлини в

90 – 95 % випадків виникають в результаті соматичних мутацій. Спадковий рак, що розвивається в результаті гермінальних змін, складає 5 – 10 % від всіх випадків захворювань [7].

Рак – це генетичне захворювання, вияв різних порушень в функціонуванні організму, як наслідок, неконтрольований поділ клітин. Мутації, що викликають онкологічні захворювання можуть успадковуватись, тоді вони є у всіх клітинах організму, або з'являються в якійсь тканині або конкретній клітині. Людина може успадкувати від батьків певну мутацію в гені, який захищає від раку, а може успадкувати мутацію яка сама по собі може спровокувати рак [8].

Для виникнення ракової клітини в ній мають статись такі події: ввімкнення онкогена і руйнація гена – супресора [9].

Мутації, що не були успадкованими, спочатку виникають в здорових клітинах. Це стається під впливом зовнішніх канцерогенних факторів. Рак щороку «молодіє», статистика захворюваності на онкологію зростає на 3 % щороку в Україні. В основному рак розвивається у людей в зрілому віці (після 50 років), після : процес виникнення та накопичення мутацій може займати більше декількох десятиків років. Нестабільність генома дозволяє пухлині збирати генетичні порушення, що можуть бути вигідними для самої пухлини. До них відносяться мутації в онкогенах – гени, які регулюють поділ клітин. Такі мутації можуть неодноразово підвищувати активність білків, робити їх нечутливими до гальмуючих сигналів чи викликати підвищене вироблення ферментів.

Це призводить до неконтрольованого поділу клітин, а з часом і до метастазування [8].

Вікові мутації – починаються з простих випадкових змін, вони накопичуються з віком в геномі будь-якої клітини організму. Сучасні дослідження показують, що деякі клітини захищені від них сильніше, наприклад, статеві клітини або нейрони мозку. Деякі – гірше захищені, такими є епітеліальні клітини. З віком таких змін стає в клітинах все більше і більше. У

цьому, мабуть, одна з головних причин того, що рак більш характерний для людей похилого віку. Коли кількість таких мутацій переходить в якість, тоді клітина стає раковою [10].

Слідом за цим сиплюються всі захисні системи клітини, а саме системи, які в ході еволюції були створені для того, щоб раковій клітині заборонити ділитися, розмножуватися і впливати на навколишні тканини і органи. Системи, які в нормі заважають її необмеженому поділу, включають гени – онкосупресори, які пригнічують розвиток раку. Клітина, яка стає раковою, позбавляється від своїх генів – онкосупресорів за допомогою генетичних змін. Часто копії таких генів просто викидаються з генома або в них вносяться мутації, які незворотнім чином блокують нормальну функцію цих генів. Тому геноми ракової і нормальної клітини можуть дуже сильно відрізнятися. Наприклад, якщо позначити послідовності кожної нормальної хромосоми у людини своїм кольором і порівняти з тим, що відбувається у ракової клітини, то в її хромосомах все буде перемішано [10].

Основні типи змін, які відбуваються в ДНК при переході від норми до раку – точкові мутації, що називаються делеціями, коли викидаються великі шматки ДНК, наприклад, брак п'яти нуклеотидів в послідовності гена BRCA2 після 999-ої «букви» призводить до раку молочної залози. Також вставки мобільних генетичних елементів – інсерції можуть спровокувати накопичення мутацій, з яких розвинеться рак. Так, у 2012 році було показано, що при більшості видів раку епітеліального походження відбуваються нові вставки мобільних генетичних елементів, що називаються генетичними паразитами, які можуть розмножувати власні копії і вставляти їх в нові місця генома, при цьому вносячи нові помилки в послідовності генів. Наприклад, такі мобільні генетичні елементи можуть вставлятися в копії генів – онкосупресорів і псувати їх, допомагаючи клітині перероджуватися ще більш ефективно. Таким чином, оскільки онкогенез характеризується не тільки масивними точковими мутаціями і делеціями, а й активацією мобільних елементів, то в майбутньому може бути зроблений ряд важливих відкриттів в галузі дослідження того, чому

при раку активуються мобільні елементи, і чи можна, блокуючи їх активність, якимись чином лікувати рак. Адже кожна вставка мобільного елемента - це серйозна зміна генома, незворотний крок клітини на шляху від нормальної до хворої. Відповідно, блокування активності мобільних елементів може допомогти ефективніше лікувати рак [10].

Мабуть, основними геномними змінами при раку є все-таки делеції або перетасовки фрагментів різних хромосом. Наприклад, коли фрагмент ДНК з однієї хромосоми переноситься на іншу і так далі (так звані генетичні транслокації). Таким чином, гени можуть потрапляти в невласне їм оточення і починають працювати інакше, зовсім не так, як вони це робили в вихідному місці, де були оточені підібраними в ході нормальної еволюції геномними регуляторними елементами. У підсумку, геном ракової клітини невідомо відрізняється від генома нормальної клітини. У порівнянні з нормою, у ракової клітини можна знайти більше функціональних копій генів, що називаються онкогенами (оскільки вони сприяють онкогенній трансформації), і менше копій генів-онкосупресорів, про які вже йшлося вище [10].

В ході придбання клітиною властивостей ракової клітини, відбувається атака на систему генів, відповідальних за репарацію ДНК. Репарація ДНК здійснює як би загоєння тих ран, які наносяться геномною ДНК мутаціями. Тобто, в нормі система репарації повинна відчувати, де саме відбуваються генетичні зміни і в ідеалі повинна ці мутації знищувати, тобто повертати генотип клітини до нормального стану. При канцерогенезі атакують саме ці системи репарації, і тим самим клітина як би позбавляється гальма на шляху подальшого накопичення мутацій, і якщо вихідна клітина мутувала зі швидкістю пішохода, то клітина з неповноцінною репарацією робить це вже зі швидкістю гоночного автомобіля і дуже швидко накопичує величезну кількість мутацій [10].

При цьому в ряді випадків, коли ракові клітини не повністю втрачають свою систему репарації, це може бути несприятливою ознакою. Це пов'язано з тим, що зараз одним з основних способів лікування деяких форм раку є

радіаційна терапія, в ході якої в геном ракових клітин вносяться численні розриви ДНК. І якщо у ракової клітини системи репарації втрачені, то вона не може знову склеїти свій геном з цих шматочків і гине. Однак, якщо якісь залишки репараційних систем у неї функціонують, то вона успішно склеює цю величезну кількість шматочків назад, і такий пазл з генів, що, певна річ, зібраний зовсім не так, як у вихідній хромосомі, продовжує функціонувати. Тобто, ракова клітина повинна, з одного боку, придушити систему репарації, але з іншого боку, для успішної протидії лікуванню, вимкнути репарацію клітина повинна в повному обсязі, залишивши її в тій мірі, щоб знову склеювати свій геном після курсів радіаційної або хіміотерапії [10].

Лікування раку в ідеалі повинно включати як би заморожку генома ракової клітини, що могло б дозволити запобігти подальшій еволюції ракових клітин в тканині. Масив клітин раку, що постійно змінюється, формує безліч генетичних варіантів, деякі з яких можуть успішно протидіяти лікуванню. Відповідно, терапія раку могла б піти по шляху фіксації генома ракових клітин в якомусь «замороженому» стані. Це могло б підвищити ефективність терапії: адже треба було б пригнічувати пухлину одного заданого типу, а не сотень тисяч різних типів, як це відбувається через мінливість ракового генома [10].

Створення детального каталогу ракових генів - важливе завдання, виконання цього завдання дозволить підбирати оптимальну терапію онкологічного захворювання для кожного пацієнта. Щоб скласти каталог ракових генів, що мутують з високою ( $> 20\%$ ) і середньої (2-20 %) частотою, для кожного гена потрібно проаналізувати в середньому 2000 пар «пухлина - норма», тобто для 50 найбільш частих типів раку це приблизно 100 000 пар. За останні 10 років вартість секвенування ДНК зменшилася в мільйон разів і зменшуватиметься ще, тому зараз це вже не являється невирішеною проблемою [11].

В організмі повинен підтримуватися баланс між активністю генів і їх продуктів, які, з одного боку, забезпечують зростання і ділення клітин, а з іншого боку – запобігають необмеженому зростанню і діленню. Зайва



активність перших або пригнічення функції других призводять до неконтрольованого росту клітин, виникнення і розвитку злоякісних неоплазій, простішими словами – ракових пухлин [11].

Гени, що пов'язані з раком можна розділити на два типи – онкогени і антионкогени (супресори пухлин), продукти яких можуть, відповідно, стимулювати або пригнічувати утворення і розвиток пухлин. Особливе місце займають мікроРНК (miRNA) – короткі (в середньому ~ 22 нуклеотида) некодуючі РНК. На сьогодні ідентифіковано більше 2000 різних мікроРНК. Вони здатні пригнічувати трансляцію мРНК, що зчитується з 30-60 % генів людини. Деякі мікроРНК сприяють злоякісній трансформації клітин, інші можуть працювати як антионкогени. У онкоген може перетворитися нормальний ген – протоонкоген, що стимулює ріст клітин постійно або на певних стадіях розвитку організму. Трансформація протоонкогена в онкоген відбувається – це результат відносно незначної модифікації його природної функції. Існують наступні основні шляхи активації протоонкогена [11]:

1) Мутація всередині протоонкогена або в його регуляторних елементах, яка змінює структуру білка і підвищує активність кодованого ним білка (ферменту) або посилює експресію відповідного гена.

2) Підвищення концентрації білка через підвищення його стабільності в клітині, збільшення періоду напіввиведення і, відповідно, підвищення активності.

3) Дуплікація гена (збільшення кількості копій), в результаті чого підвищується концентрація білка в клітині.

4) Транслокація гена, що викликає посилення його експресії або виникнення агресивного гібридного гена.

Функція антионкогенів протилежна функції протоонкогенів. Антионкогени контролюють різні процеси, що перешкоджають злоякісній трансформації клітин [11]:

- 1) Придушення надмірної експресії генів, які забезпечують проліферацію клітин.
- 2) Здійснення репарації ДНК (пошкодження ДНК при придушенні репарації підсилюють мутагенез і, як наслідок, активацію протоонкогенів і інактивацію антионкогенів).
- 3) Координація проліферації клітин з репарацією ДНК. Якщо репарація ДНК пригнічена, вони гальмують поділ клітин і ініціюють апоптоз.
- 4) Контроль адгезії(здатність клітин формувати певні типи гістологічних структур) і механізмів контактного гальмування клітин, які діляться [11].

Загалом, антионкогени ставлять бар'єр необмеженого росту клітин. Втрата функції антионкогена руйнує цей бар'єр. Найвідомішим і часто мутуючим в ракових пухлинах багатьох типів є антионкоген TP53. Продукт TP53 – фосфопротейн регулює транскрипцію ряду різних генів. У нормальній клітині він неактивний. При надзвичайній події він активується і виконує роль «охоронця геному», здійснюючи різні протиракові функції [11]:

- 1) Активація системи репарації ДНК.
- 2) Якщо ДНК пошкоджена, TP53 затримує мітоз клітин, які діляться, надаючи системі репарації час усунути пошкодження.
- 3) Якщо усунути пошкодження ДНК не вдається, TP53 включає програму загибелі клітин – апоптоз.

У разі, якщо рак викликаний протоонкогеном, зазвичай досить активації цього протоонкогена на одній з двох парних клітинних хромосом. Але якщо рак виник через втрату ефекту антионкогена, то, як правило, потрібні мутації або втрата обох його копій [11].

### 1.3 Порівняльне секвенування

На сьогодні ідентифіковано приблизно 300 онкогенів і супресорів пухлин. Ракові гени шукають використовуючи порівняльне секвенування – порівнюють послідовності нуклеотидів в ДНК пухлин і нормальних тканин і потім ідентифікують відсутні в ДНК нормальної тканини соматичні мутації, такі мутації зустрічаються частіше, ніж просто випадкові події [11].

Ця стратегія реалізується в декілька етапів. В першу чергу, потрібно отримати зразки пухлинної тканини від пацієнтів з достовірно поставленим діагнозом і детально описаною картиною протікання захворювання. При цьому зразки повинні бути, по можливості, вільні від нормальних клітин. Для порівняння використовуються проби крові пацієнта або нормальної тканини. З тканини пухлини і нормальної тканини виділяють ДНК і досліджують їх за допомогою секвенування. Останнім часом секвенування проводиться за допомогою платформ нового покоління, що дозволяють швидко і порівняно недорого повністю секвенувати геном людини. Результати порівняльного секвенування потім аналізуються за допомогою спеціально розроблених дуже складних математичних і біоінформатичних методик (рис. 1.3) [11].

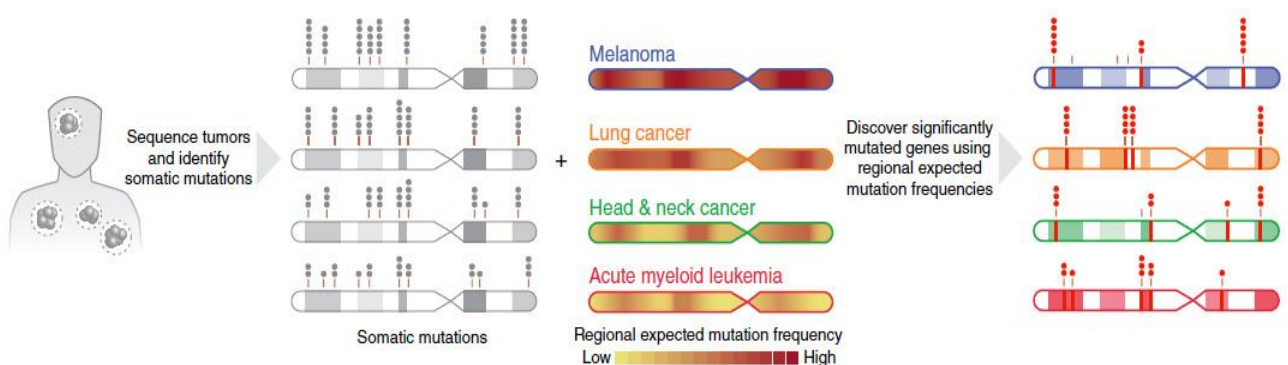


Рисунок 1.3 – Результати порівняння послідовностей нуклеотидів [11]

На цілісність геному впливають такі зовнішні фактори: вірусні інфекції, куріння, ультрафіолетове випромінювання. Різні ділянки генома мутують з різною частотою в залежності від форми раку. Результати порівняння

послідовності нуклеотидів в ДНК пухлин і нормальних тканин аналізуються за допомогою математичних і біоінформатичних методик, що враховують частоти і спектр мутацій, порядок, в якому відбувається реплікація сегментів ДНК по всій довжині хромосоми і рівень експресії генів також. Такий метод дозволяє знайти гени, достовірно мутуючі при раку [11].

Фінальною метою цих проектів є створення детального каталогу аномалій структури генома, які пов'язані з ініціацією, прогресією і підтриманням онкологічних новоутворень. Такий каталог дозволить отримати багато нових даних по молекулярної біології раку та удосконалити методи діагностики, лікування та профілактики раку, визначати нові мішені для розробки протиракових засобів. Систематичні дослідження за цим напрямком вже дозволили ідентифікувати не тільки багато нових ракових генів, а й навіть цілі класи ракових генів [11].

#### 1.4 Помірні і рідкі онкогени

Деякі пов'язані з раком гени мутують досить часто - при багатьох або, хоча б, при декількох типах раку. Тому не дивно, що саме вони (зокрема, TP53) були охарактеризовані першими [11, 13].

Мутація в гені TP53 – необхідна умова летального раку. У 55 % випадків раку в людини у ракових клітинах виявляється зміна гену TP53, а у випадку раку легенів мутація виявляється більш, ніж у 90 %. TP53 – ген, що пригнічує ракові клітини. Мутація в даному гені завжди відбувається останньою [11].

Більшість ракових генів виявляються з помірною частотою (2-20%) або рідше. Основні проблеми виникають при ідентифікації рідкісних ракових генів. Так, недавнє дослідження 183 аденокарцином легенів показало, що у 15% пацієнтів не виявляється жодної мутації в 10 відомих для цього захворювання класах генів, а в 38% випадків ідентифіковано три або менше мутацій [11, 13].

Внаслідок пошкодження систем репарації ДНК в ракових клітинах в певній мірі посилюється мутагенез. Частота цих індукованих пухлиною соматичних мутацій для різних пухлин може відрізнятися на кілька порядків. Тому при ідентифікації «помірних» і особливо «рідкісних» генів виникає істотна проблема: як відрізнити гени і мутації, що пов'язані з раком, від фону, від численних випадкових мутацій, що не мають відношення до раку [11]?

Гарвардська група під керівництвом Г. Гетца розробила програмний пакет MutSig (скор. Від 'Mutation Significance'), що аналізує частоти і спектр мутацій в різних областях генома при різних формах раку, а також ряд інших факторів, пакет дозволяє виокремити гени, що достовірно мутують при раку [11].

### 1.5 Гени достворіно мутуючі при раку

Було проведено дослідження, де вивчили зібрані з різних баз даних результати аналізу екзомів (кодуючих областей генів – екзотів з прилеглими до них послідовностями ДНК – некодуючих проміжків – інтронів) 4742 пар «пухлина – нормальна тканина», що належать до раку 21 різних типів. Кількість зразків для окремих типів варіювало від 35 до 892 [11].

Було виявлено 3 078 483 поодинокі заміни нуклеотидів в пухлинах в порівнянні з нормальною тканиною, 77 270 одиничних делецій або вставок нуклеотидів, 29 837 ди-, три- або олігонуклеотидних делецій або вставок. З поодиноких замін нуклеотидів переважна більшість (2 294 935) не змінювало кодуючі послідовності. З інших одиничних замін 540 831 представляли собою так звані місен – мутації (що призводять до замін амінокислот в білках), 207 144 представляли собою синонімічні (НЕ призводячі до замін амінокислот) заміни нуклеотидів, 46 264 – нонсенс – мутації (що призводять до дострокової термінації синтезу білка), 33 673 – мутації, що порушують сплайсинг мРНК

(збірку кодуєчих послідовностей мРНК з відповідних блоків). Дані по «глибині» секвенування і чистоті зразків пухлин дозволяють оцінити чутливість аналізу величиною більше 90 % (рис. 1.4). Частоти виникнення мутацій на одиницю довжини геному для різних типів раку відрізнялись більше ніж на 5 порядків (від 0,03 до 7000 на мільйон нуклеотидів ДНК), сильно розрізнялись і спектри мутацій [11].

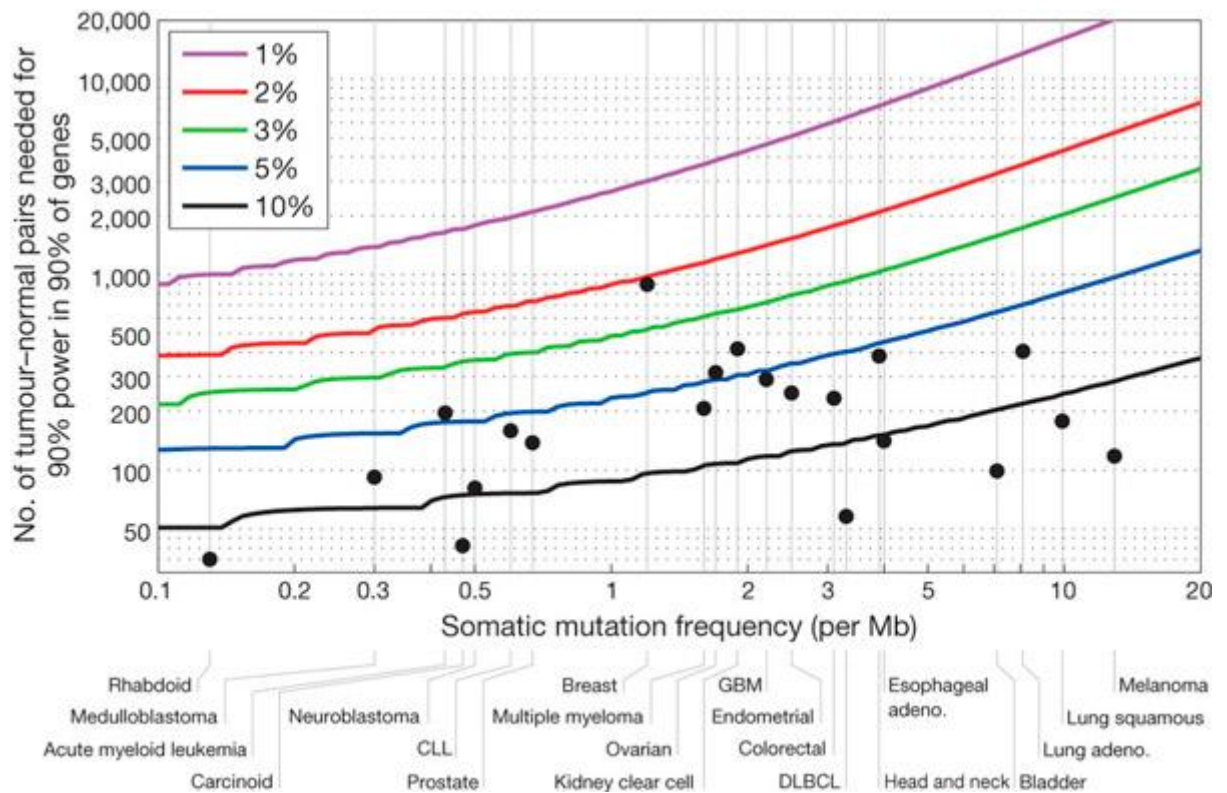


Рисунок 1.4 – Кількість пар «пухлина-норма» (по осі ординат), необхідних для визначення 90% генів, достовірно мутуючих при раку з вірогідністю 90 % [11]

Оцінки приведені в залежності від того, типу пухлини, середньої частоти фонових мутацій (по осі абсцис), перевищення частоти мутацій ракових генів над фоном (кольорові лінії). Для більшості типів пухлин ще недостатньо кількості проаналізованих зразків навіть для того, щоб детектувати гени, що мутують з частотою над фоном 5 % и менше. Чорними точками показано вже проаналізовану кількість зразків [11, 14].

Мутації, достовірно пов'язані з раком, виявлялися в 224 різних генах. Для різних типів раку число мutowаних генів сильно варіював (від 1 до 58). Для 7 типів воно було менше 10, а для двох типів, такі як рак молочної залози і ендометрія матки, понад 30. Лише 22 гена виявилися достовірно пов'язаними з більш ніж трьома типами раку. Аналіз дозволив ідентифікувати майже всі раніше відомі гени, пов'язані з канцерогенезом. Знайдено також 33 гена, мутації, які раніше не порівнювались з раком. Ці гени пов'язані з поділом клітин, апоптозом, стабільністю генома, імунною відповіддю, регуляцією активності хроматину, гомеостазом білків і перетвореннями РНК. Серед ще 81 гена також повинні бути гени, пов'язані з раком [11].

На підставі отриманих результатів було зроблено розрахунки такі як: скільки пар «пухлина - нормальна тканина» потрібно проаналізувати, щоб знайти точно мутуючі при раку гени в залежності від типу раку, частоти мутацій на одиницю довжини геному при певному типі, частоти мутацій даного гена при цьому типі раку. Проведені розрахунки показують, що для 17 з 21 проаналізованого типу раку даних ще недостатньо, щоб виявити гени, що мутують з частотою не вище 5 % від фону. А для 7 типів - навіть при частоті не вище 10 %. Для складання каталогу ракових генів, що перекривають 90 % випадків захворювання, необхідно проаналізувати близько 650 пар «пухлина - норма», якщо середня частота мутацій становить близько 0,5 на мільйон пар основ ДНК (як при нейробластомі). Або навіть близько 5300 пар, якщо частота десь 12,9 на мільйон (як при меланомі). Щоб скласти каталог соматичних мутацій для генів, мутуючих з високою (> 20 %) і середньої (2-20 %) частотою для вже відомих біля 50 типів раку, потрібно проаналізувати в середньому 2000 пар «пухлина - норма», тобто всього приблизно 100 000 пар [11].

## 1.6 Рак стравоходу

Розглянемо накопичення мутацій в стравоході під час старіння. В нормальному епітелії стравоходу протягом життя накопичуються мутації, але вони не обов'язково викликають рак [15].

Візьмемо дві групи людей, детальний аналіз структури генів, що вважається асоційованим з раком епітелію стравоходу, у різному віці. Було виявлено, що в результаті природних помилок при реплікації ДНК в тканинах стравоходу з віком виникає все більше мутацій і навіть у здорових людей більшість помилок були накопичені в «онкогенних» генах. Злоякісного переродження не сталося через те, що набори мутацій в здорових тканинах сильно відрізняються від наборів, характерних для ракових захворювань [15].

При кожному поділі ДНК материнської клітини передається дочірнім клітинам. В ідеалі кожна з них має отримати по одній точній копії всіх молекул ДНК, які були у материнській клітині. Подвоєння молекул ДНК відбувається в процесі реплікації. Проблема в тому, що системи реплікації кліткового геному можуть інколи помилятися, тобто змінювати послідовність нуклеотидів ДНК [15].

Вчені секвенували 74 гени, які частіше всього асоціюються з різними типами раку (тобто гени, зв'язок яких з онкогенезом або доказано або підозрюється). У вивчених зразках від молодих людей спостерігалось набагато менше мутацій, чим у зразках від людей похилого віку [15].

З перевірених потенціальних онкогенів самим часто мутуючим в здоровому епітелії виявився NOTCH1. Цей ген кодує один з трансмембранних рецепторних білків і асоціюється з деякими видами раку, в тому числі і плоскоклітинною карциномою стравоходу (oesophageal squamous cell carcinoma, OSCC) – самим поширеним видом раку в Південній Америці і Азії. Показано, що одні з основних факторів ризику для нього – вживання алкоголю і куріння. Мутації цього гена спостерігаються приблизно у 10% випадках



плоскоклітинної карциноми стравоходу. У вивчених зразках епітелію стравоходу спостерігалось в середньому 120 різних мутацій NOTCH1 в перерахунку на квадратний сантиметр, де Synonymous- синонімічні мутації. Missense – місенс – мутації, Nonsense – нонсенс – мутація, термінуючі трансляцію мРНК, Splice – порушення сплайсинга, Indels – інші мутації (рис. 1.5) [15]. У людей похилого віку частота мутацій в цьому гені була помітно вище ,ніж у молодих. Але ось,що важливо: було знайдено порівняно мало супутніх мутацій в гені TP53(які були виявлено більше,ніж в 90 % пухлин OSCC): їх було приблизно в 10 разів менше,ніж в середньому в клітинах ракових пухлин. Також було знайдено порівняно мало аномалій структур хромосом( це признак злоякісних клітин , типових для клітин OSCC і багатьох інших раків) [15].

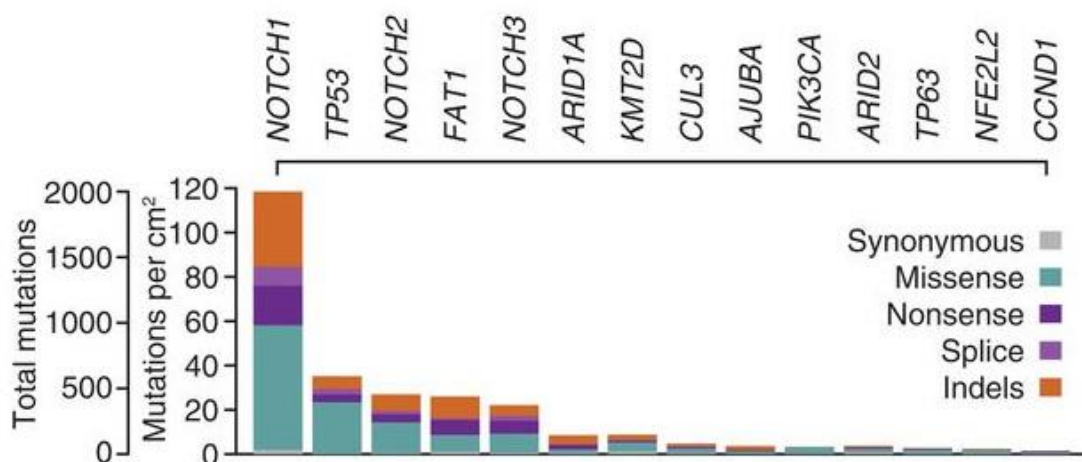


Рисунок 1.5 – Частота мутацій в 14 генах, що асоціюються з раком в епітелії стравоходу [15]

В здоровому епітелії найчастіше мутує ген NOTCH1, а в пухлинах - TP53.

Загалом отримані в цій роботі результати показали, що в здоровому стравоході мутації накопичуються головним чином внаслідок нормальних фізіологічних процесів, помилок при реплікації ДНК і що вони зовсім не обов'язково призводять до раку. Щоб розвинувся рак, цей процес повинен зайти далі: відбуваються додаткові мутації, порушується контроль реплікації генома,

виникають хромосомні аномалії. Цьому може сприяти додатковий вплив зовнішніх канцерогенів [15].

Нові мутації постійно виникають, більшість з них нейтральні - не шкідливі і не корисні. Найчастіше мутації відбуваються в генах, асоційованих з раком, хоча при цьому зовсім не обов'язково, що рак розвивається. Клітини з такими мутаціями залишаються нормальними, але можуть отримати переваги перед сусідами, в результаті чого їх клони можуть заповнити значну ділянку стінки стравоходу. Експансія клонів мутантних клітин епітелію стравоходу сама по собі раку не викликає, але без неї, судячи з усього, ракова пухлина не розвинеться - для цього повинні відбутися ще якісь події, наприклад, додаткові мутації, що ослаблюють захисні механізми клітини [15].

### 1.7 Рак молочної залози

При раку молочної залози основні мутації здійснюють накопичення ще в передраковому стані.

Рак молочної залози – саме поширене онкологічне захворювання у жінок. Майже у 10 % жінок рано чи пізно розвивається рак цього типу. Чоловіки теж зіштовхуються з таким видом онкології, але дуже рідко. Як правило рак молочної залози – це результат злоякісного переродження клітин протоків молочної залози або долей, в яких виробляється молоко, всього відомо 18 форм цього раку [16].

Найчастіша і добре вивчена форма раку молочної залози (на неї припадає близько 80 % всіх випадків) – дуктальна карцинома. У багатьох випадках, хоча і не завжди, раку передуює карцинома проток молочної залози *in situ* (ductal carcinoma *in situ*, DCIS). Це вважається або передраковим станом, або ранньою стадією раку. Вона зазвичай виявляється при обстеженні молочної залози стандартною процедурою – мамографією. Передраковим станом DCIS

називають тому, що вона являє собою кластер злоякісних клітин, що не є схильними до інвазії в сусідні тканини [16].

З ймовірністю в межах 10-30 % DCIS перетворюється в справжній рак, інвазивну карциному проток (invasive ductal carcinoma, IDC), клітини якої вже можуть поширитись в навколишні тканини і утворювати метастази. DCIS може перебувати в латентному стані багато років. Точно передбачити варіант «пробудження» DCIS і її перетворення в інвазивну карциному неможливо, тому хворим рекомендується превентивна хірургічна операція з наступною променевою терапією [16].

Раніше наводилось два сценарії цього перетворення. Згідно з першим варіантом, DCIS і IDC паралельно розвиваються з нормальних клітин, які незалежно один від одного перетворюються в злоякісні з різноманітними властивостями. Цей сценарій підкріплюється дослідженнями, де були виявлені різні молекулярні маркери у DCIS і IDC. Згідно з другим варіантом, в DCIS формуються різні популяції клонів злоякісних клітин, але тільки одна з них проходить через «пляшкове горлечко», іншими словами долає базальну мембрану і дає початок IDC. На користь цього сценарію говорять геномні дослідження, де було виявлено однакові мутації в клітинах DCIS і IDC. Але досі невідомо як саме відбувається перетворення карциноми проток *in situ* в інвазивну карциному [16].

У DSIC міститься порівняно мало клітин, а IDC гетерогенна (мутації в її клітинах занадто різноманітні), тому існуючі методи для порівняння властивостей DCIS і IDC не підходили. Автори роботи модифікували, вдосконалили і звели воедино ряд всі відомих методи. В результаті вийшов новий підхід, що називається топографічним секвенуванням окремих клітин (Topographic Single Cell Sequencing, TSCS). З TSCS вдалося простежити, які зміни відбуваються в геномах окремих клітин при виникненні DCIS з нормальних тканин і при перетворенні DCIS в IDC [16].

Вчені брали зразки тканини у пацієнтів з DCIS і IDC і на гістологічних зрізах, пофарбованих гематоксилін – еозином (рис. 1.6 А), визначали місце

знаходження клітин цих захворювань за морфологічними відмінностям від нормальних клітин. Також визначалися області з нормальною тканиною. Потім з цих ділянок лазером вирізали окремі клітини і переміщали їх в окремі пробірки теж за допомогою лазера (рис. 1.6 В, С). Такий підхід дозволив звести до мінімуму ймовірність забруднення стороннім матеріалом. Після цього вчені почали геномні дослідження, підготовка до дослідження показана на (рис. 1.6 D). Клітини лізували і ампліфікували їх ДНК за допомогою полімеразної ланцюгової реакції з «виродженими» праймерами [16].

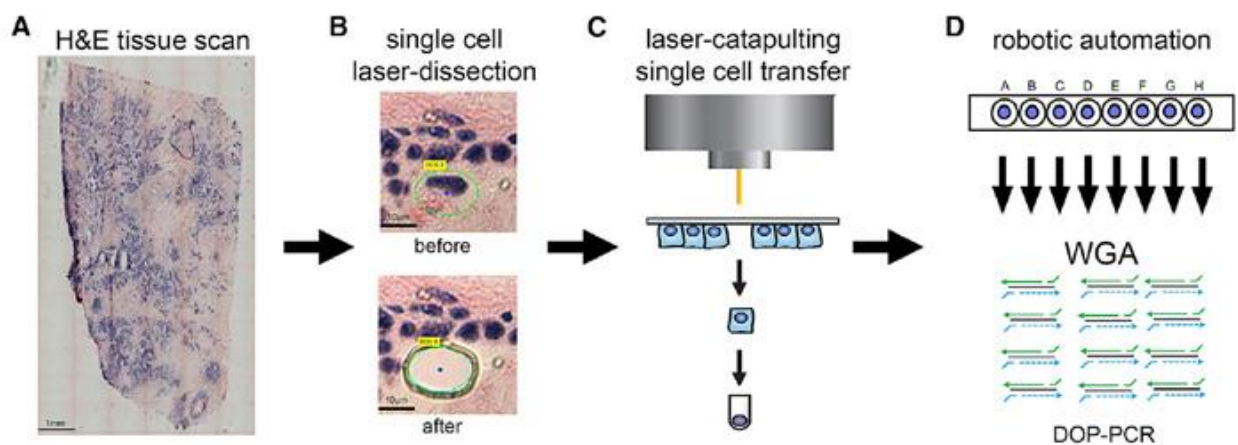


Рисунок 1.6 – Схема топографічного секвенування окремих клітин [16]

Отриманий матеріал аналізували методами високоефективного секвенування нового покоління. Таким чином були охарактеризовані 1293 пухлинні клітини, які отримані від 10 хворих на рак молочної залози. Високоефективне секвенування дозволяє не лише визначити послідовність нуклеотидів ДНК, а й оцінити кількість копій різних ділянок ДНК в геномі. В результаті з'ясувалось, що і в клітинах DCIS, і в клітинах IDC суттєво варіює число копій генів [16].

## 1.8 Висновок до розділу 1

Важливо знати основні поняття в будові ДНК та класифікацію мутацій, для подальшого розуміння та аналізу методів для виявлення генетичних змін. Як ми бачимо, з вище описаного, рак – процес неконтрольованого поділу клітин, мутації, що викликають онкологічні захворювання.

Ефективність хіміотерапії є наслідком випадкового збігу і є найбільш токсичним видом лікування, на пізніх стадіях ракові клітини перестають реагувати на терапію. Тому рання діагностика є вкрай важливою для успішного лікування. Чим більша пухлина, тим вища ймовірність наступної мутації. Такі зміни важливо знайти і розпізнати в послідовності ДНК. Секвенування направлене на вирішення даної задачі. Методом порівняльного секвенування і аналізу послідовностей нуклеотидів при нормі і патології було знайдено і класифіковано ракові гени. Тому дана робота націлена на вивчення методів секвенування нового покоління, а саме нанопорове секвенування, аналіз сигналів і класифікацію вище описаних мутацій.

## 2 МЕТОД НОВОГО ПОКОЛІННЯ – НАНОПОРОВЕ СЕКВЕНУВАННЯ ДНК

Секвенування (з англ. sequence – «послідовність») – це загальна назва методів, за допомогою яких встановлюють послідовність нуклеотидів в молекулі ДНК [17].

Oxford nanopore Technologies-компанія у Великобританії, яка займається розробкою та продажем продуктів секвенування нанопор для прямого електронного аналізу одиничних молекул [18].

Нанопорове секвенування (Nanopore sequencing) відноситься до NGS (секвенування нового покоління) і являє собою наступний етап розвитку технологій секвенування нуклеїнових кислот [19].

При нанопоровому секвенуванні відбувається читання нуклеотидного складу безпосередньо молекули, що аналізується, яке приводить до спрощення процесу, зменшення часових затрат (в першу чергу, на аналіз результатів) і зниження його вартості. Принцип нанопорового секвенування заключається в переміщенні молекули, що аналізується, через пору і реєстрації змін в оточуючому електричному полі в процесі руху молекули [19]. Всі попередні технології ґрунтуються на непрямому визначенні послідовності, тобто включає в себе стадію синтезу на ДНК-матриці, в процесі якого відбувається детекція вбудованих нуклеотидів-мономерів. При цьому процес секвенування включає стадію пробо підготовки (ПЛР – полімеразна ланцюгова реакція, мічення зразка і т.д.) [19]. В теперішній момент для секвенування використовуються нанопорові структури біологічного походження, але при цьому розробляються штучні нанопори на основі діоксиду кремнія, а також гібридні молекули [19].

Нанопора представляє собою отвір в нано-маштабі. В приладах Oxford Nanopore через нанопори пропускається іонний струм і прилад вимірює зміну струму, коли молекули проходять через нанопору. Інформацію про зміну струму можна використовувати для ідентифікації зміни біомолекул. Такі

отвори можуть бути утворені білками, пронизуючими мембрани (біологічні нанопори) або отворами в твердих матеріалах (так звані, твердотільні нанопори) [20].

## 2.1 Біологічні нанопори

В технологіях першого покоління Oxford Nanopore використовуються індивідуальні пороутворюючі білки, що створюють пори в мембранах. Різноманітні пороутворюючі білки являються дуже схожими по своїй природі. Наприклад, білок  $\alpha$ -гемолізін і подібні йому пороутворюючі білки виявляються в природі і клітинних мембранах, де вони функціонують в якості каналів для переносу іонів і молекул як в клітину, так і в навколишнє середовище.  $\alpha$ -гемолізін представляє собою гептамірну білкову пору з внутрішнім діаметром 1 нм. Це приблизно в 100 000 раз тонше людської волосини. Цей діаметр має той самий розмір як і багато одиничних молекул, включаючи ДНК. Така пора дуже стабільна [20].

## 2.2 Як працюють нанопори при секвенуванні ДНК

Нанопорові системи представляють собою реакційну камеру, всередині якої знаходиться розчин електроліту. Камера розділена на дві частини ліпідною мембраною або іншою тонкою непровідною поверхнею, в яку поміщають одиничну нанопору (рис. 2.1). До частин камери прикладають напругу, через що виникає струм іонів через пору. Коли досліджувані молекули проходять через пору за напрямком дії поля, вони зменшують переріз, доступний для

іонів, і сила струму падає. Аналізуючи зміни сили струму, можна визначити властивість молекули, що проходять через пору (рис.1.1) [21].

У випадку нуклеїнових кислот, діаметр використовуваних нанопор складає декілька нанометрів, через що ДНК і РНК здатні проходити через пору тільки в одноланцюговій формі, але не в дволанцюговій. При проходженні молекули нуклеїнової кислоти через пору окремі нуклеотиди затримуються в певних сайтах всередині пори, не даючи пройти іншим іонам, в результаті чого відбувається вимірюване падіння струму [21, 22]. Струм через пору пропорційний розміру нуклеотиду, що проходить [23].

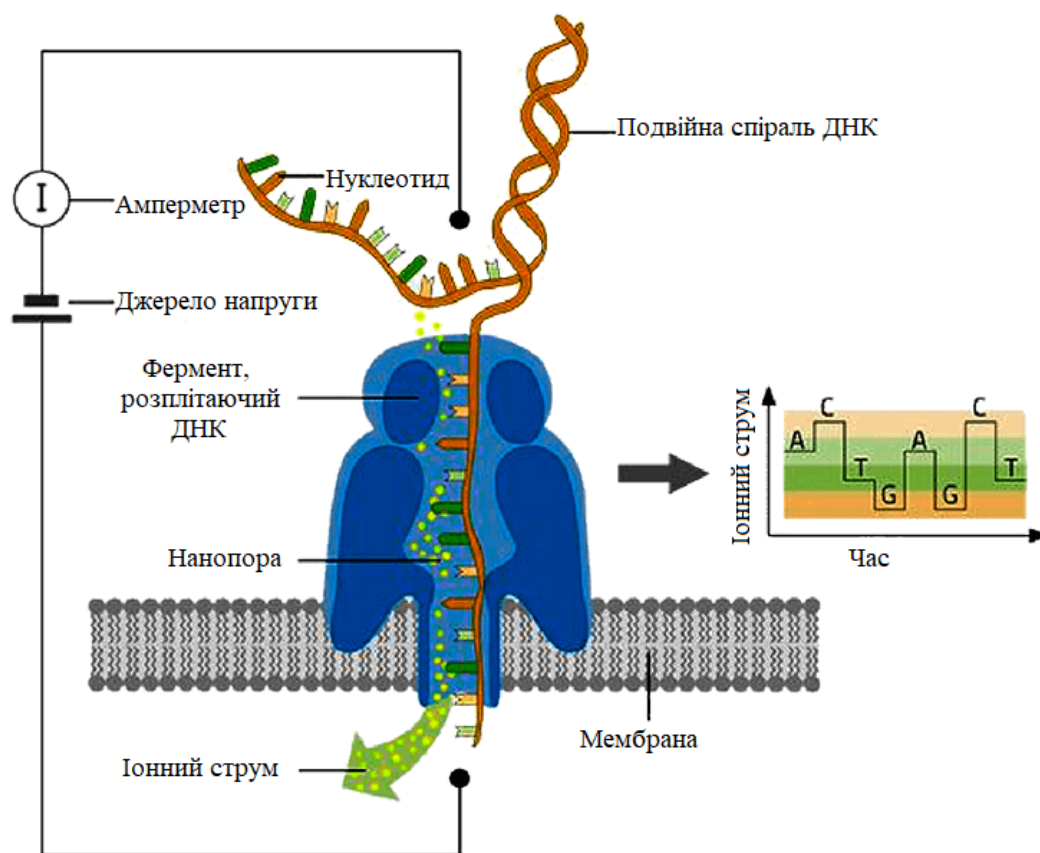


Рисунок 2.1 – Білкова нанопора, закріплена в бішаровій мембрані



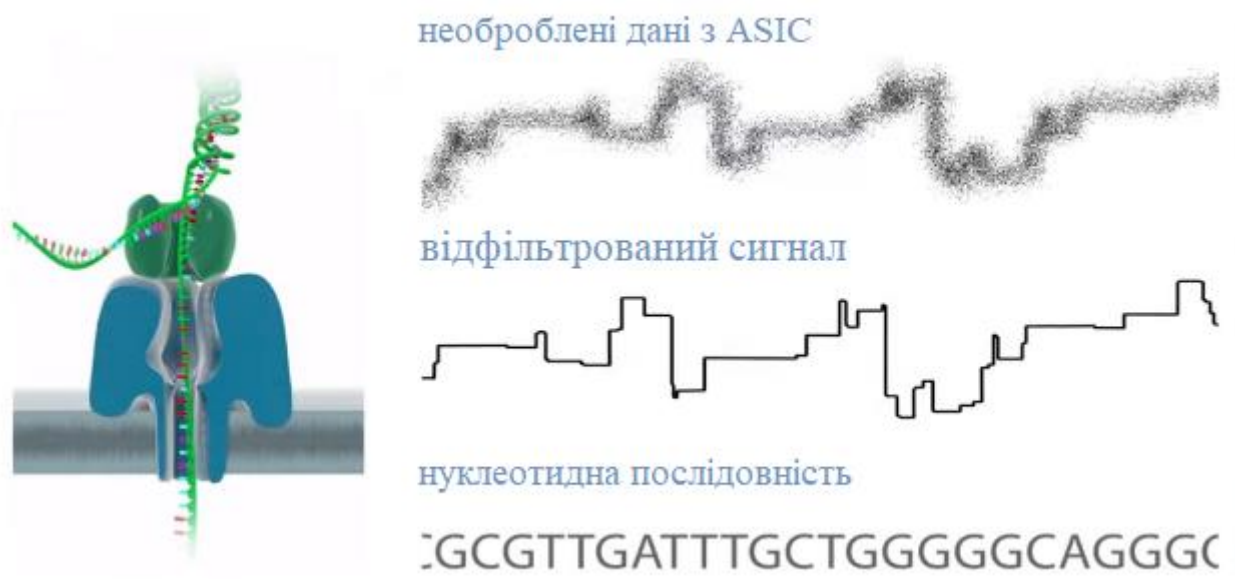
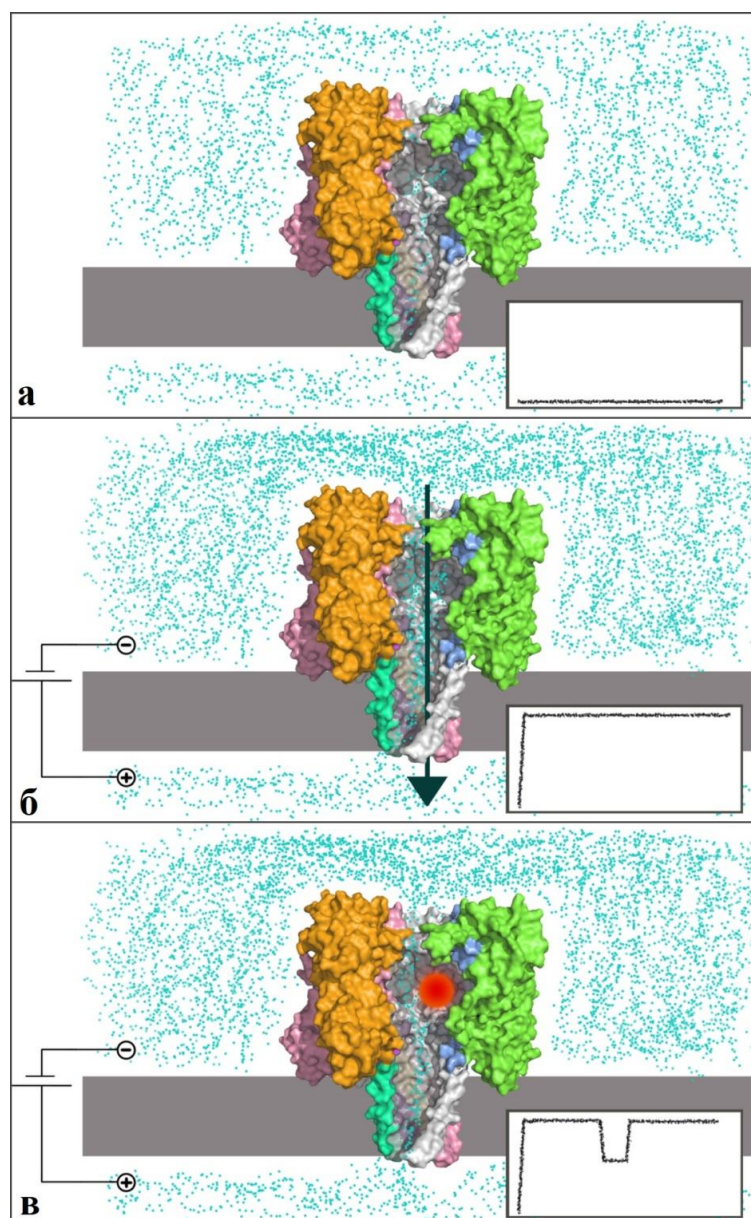


Рисунок 2.2 – Процес секвенування всередині пори

Молекула ДНК або РНК, несуча на собі суттєвий негативний заряд, під дією електричного поля «протискується» через білкову пору, поміщену в бішарову мембрану, що викликає зменшення сили електричного струму через зменшення перерізу отвору (Рис. 2.2). Зчитуючи зміну струму, нанопоровий секвенатор визначає нуклеотид, що проходить через пору в конкретний відрізок часу [21].

Зміна струму іонів через канал при проходженні нуклеотида складає до 40 пА і помітно відрізняється для кожного з чотирьох типів складаючих спадкову молекулу нуклеотидів. Ці параметри успішно визначають нуклеотидну послідовність ДНК [24].



а – напруга відсутня; б – під дією прикладеної напруги виникає струм іонів; в – при проходженні великої молекули через пору струм іонів зменшується.

Рисунок 2.3 – Нанопора  $\alpha$ -гемолізіну в мембрані

## 2.3 Варіанти нанопорового секвенування

### 2.3.2 Секвенування цілих ланцюгів

В даному методі ланцю нуклеїнових кислот не розщиплюються. Перенесення цілих молекул ДНК і РНК через пору може здійснюватися наступними способами:

### **Транспорт під дією напруги**

Так як ДНК і РНК несуть на собі негативний заряд, то найпростішим способом транспорту молекули нуклеїнових кислот через пору являється електрофоретичний перенос разом з іонами. Проблемою даного методу являється те, що для виміру падіння струму іонів через пору з самого початку потрібен великий струм, щоб отримати хороше співвідношення сигнал/шум. Але при збільшенні прикладеної напруги збільшується швидкість, з якою молекула нуклеїнової кислоти долає пору, а значить зменшується час розпізнавання кожної окремої основи, через що якість розпізнавання падає [25].

### **Транспорт під дією напруги з розплітанням дуплексів**

Зменшити швидкість проходження одно ланцюгової ДНК через пору можна, утворюючи з нею дволанцюгової ділянки з допомогою комплементарних фрагментів ДНК. Тоді в ході транспорту буде відбуватись розплітання даної ділянки, що і дозволить довше затримувати окремі нуклеотиди в порі. Тим не менш, оскільки розплітання відбувається не по нуклеотидно, то час затримки нуклеотида в порі не являється постійним для всієї послідовності [25].

### **Транспорт з використанням ферментів**

Для того, щоб кожний нуклеотид затримувався в порі на фіксований час, можна використовувати різноманітні ферменти, які будуть пропускати нуклеотиди через пору по одному. Прикладом такого ферменту є ДНК – полімераза. За рахунок прикладеної напруги комплекс ДНК – фермент з самого початку притягнутий до пори. Але тепер, перш, ніж черговий нуклеотид молекули ДНК пройде через пору, повинен статися ще один крок синтезу другого ланцюга ДНК. Утворююча затримка основи всередині пори дозволяє більш точно розрізняти їх [25].

## 2.4 Екзонуклеазне секвенування

В даному методі ланцюг нуклеїнової кислоти нарізається на одиничні нуклеотиди нуклеазою, розміщеною в безпосередній близькості від пори. Під дією поля негативно зарядженні нуклеотиди самотійно потрапляють в пору, де відбувається визначення основ [25].

## 2.5 Типи нанопор

Розрізняють два типа білкових нанопор: білкові нанопори (альфа – гемолізін, MspA, моторний білок phi29), твердотільні нанопори (флуоресцентна детекція).

### 2.5.1 Білкові нанопори

#### **Альфа-гемолізін:**

В арсеналі нанотехнологів є перевірена і запатентована нанопора, в структуру якої входить гарно вивчений білок альфа-гемолізін. Цей бактеріальний протеїн вбудовується в клітинну мембрану і дірявить її, формуючи пору [26].

Альфа-гемолізін (білок із стрептококу) — водорозчинний мономер, який в мембрані самовільно утворює гептамер. Альфа-гемолізін служить для проходження іонів, фактично іонний канал. ДНК може проходити через цю пору точно так само як і іон [26].

Іонні канали — інструмент, через який клітина здійснює обмін речовин з зовнішнім світом. Ці складні молекулярні пристрої одночасно аналізують молекули і контролюють потік частинок крізь клітинну мембрану. Німецькими вченими було створено штучну версію такого каналу, їх аналог може виконувати обидві функції. Пропускна здатність іонних каналів визначають наступні параметри: форма, діаметр, заряд, що проходить через канал молекули), і цими параметрами можна варіювати. Аналітичні можливості штучних каналів, як було підтверджено рядом експериментів, теж можна порівняти з природними, оскільки вони надійно розрізняють полімерні молекули по конфігурації і довжині [27].

Альфа-гемолізін складається зі стовбура і головки пори. Головка прикріплюється до мембрани гідрофобним донцем. До донця приєднується стовбур з залишків лізину і глутамата, він і проходить через мембрану [27].

Головка пори містить порожнину діаметром близько 4,5 нм. У місці з'єднання стовбура і головки знаходиться звуження каналу шириною 1,5 нм. Стовбур пори складається з 14 антипаралельних бета-шарів, що формують наскрізний канал шириною близько 2 нм. В середині стовбура знаходиться три сайти розпізнавання нуклеотидів, що в теорії робить можливим розпізнавання одного нуклеотида більш, ніж одним сайтом (що збільшить точність прочитання). Час перебування різних нуклеотидів на цих сайтах різниться, що і дозволяє відрізнити одні основи від інших [25].

Вивчивши властивості шкідливих гемолізінів, вчені загальними зусиллями привчили білок, змусивши його секвенувати нуклеотиди. Проте структура гемолізинової нанопори створює деякі обмеження для ефективного аналізу ДНК. Так, через велику глибину каналу (5 нм) специфічність нуклеотидів, що виражена напругою іонного струму, зменшується. Тобто сигнал розмивається і можна сплутати нуклеотиди [26].

### MspA:

Нанопора діаметром 1,2 нм у вигляді воронки. Володіє структурними особливостями ( формою і діаметром пори), які покращують співвідношення сигнал/шум при секвенуванні ДНК порівняно з альфа-гемолізином (рис. 2.4) [25].

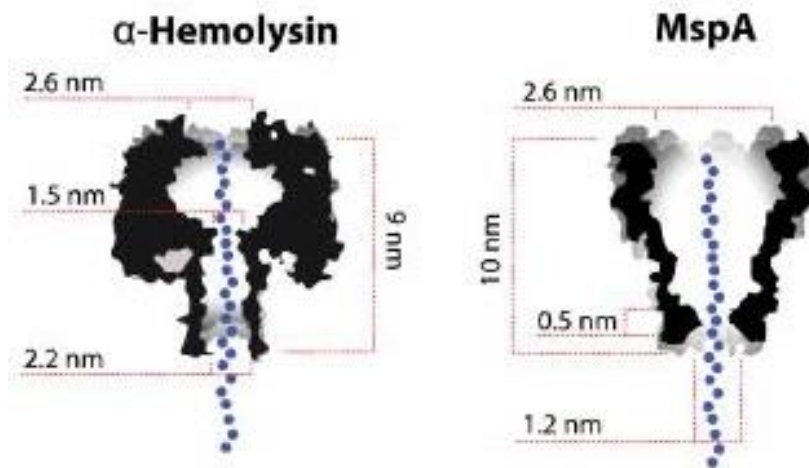


Рисунок 2.4 – Біологічні пори  $\alpha$ -гемолізину і MspA

В центрі білкової воронки створюється канал глибиною всього 2 нм. Звуження у 1,2 нм дозволяє пройти крізь воронку всього лише одному нуклеотиду. [26]

На просторовій моделі (рис. 2.5), кольорами відмічено розподіл заряду всередині. Червоним і синім кольорами позначені позитивно і негативно заряджені ділянки відповідно; фіолетові ділянки мають подвійний заряд (полярні основи)); жовті та помаранчеві – гідрофобні основи (парафінові і ароматичні) [28].

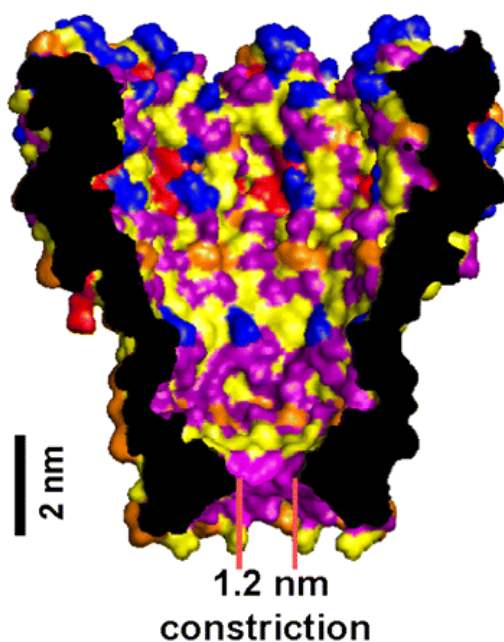


Рисунок 2.5 – 3D – модель нанопори в поперечному розрізі [28]

Дослідники припустили, що MspA порин зможе розпізнавати нуклеотиди. MspA порин має надлишковий негативний заряд, який також може створювати перешкоди, тому експериментатори з допомогою генетичних маніпуляцій нейтралізували носіїв надлишкового заряду — три основи аспарагінової кислоти [26].

#### **Моторний білок phi29:**

Моторний білок з'єднується з порою і з використанням АТФ проштовхує через пору частину ДНК під дією електричного струму.

У бактерій існує ротаційний мотор, що нагадує за своєю структурою електродвигун. Він використовується для переміщення клітин в водному середовищі. Спеціальні білки розвивають значне механічне підсилення і здійснюють переміщення при роботі з молекулами ДНК [29].

Часто стояла проблема в пропусканні ланцюга ДНК крізь канал з більш-менш постійною швидкістю — достатньо швидко для того, щоб такий метод секвенування мав сенс, але при цьому, щоб іонний струм встигав змінюватись і ці зміни можна було зафіксувати. Вирішити цю задачу вдалось, приєднавши до порину ДНК-полімеразу фага (бактеріального вірусу) phi29 — генетичного

паразиту бактерій *Bacillus subtilis*. Цей компонент забезпечив рівномірне і з потрібною швидкістю протискування ДНК через канал [30].

Недоліком біологічних моторів являються спеціальні умови: рідке середовище певного сольового розчину, температура і рН. Це обмежує область їх застосування [29].

### 2.5.2 Твердотільні нанопори

Окрім білкових нанопор також використовуються твердотільні нанопори небіологічної природи. Для аналізу нуклеїнових кислот використовуються нанопори в підкладках із кремнію, нітриду кремнія і поліетиленіміну. Зазвичай пори формуються пучками іонів або електронів, що дозволяє легко варіювати їх розмір. Окремо варто виділити матеріали, що формують дуже тонкі — «2D»-пори: графен, дисульфід мольбдена та інші. Графен володіє і достатньо маленькою товщиною, що сприяє збільшенню простору вздовж ДНК, і одночасно з цим міцністю, електропровідністю. Ці властивості роблять його дуже перспективним матеріалом в нанопоровому секвенуванні [25, 31].

#### **Флуоресцентна детекція:**

В 2010 році був розроблений метод твердотільного наносеквенування, заснований на детекції флуоресцентного сигналу. Для початку потрібна ДНК конвертується в ДНК, в якій кожній початковій основі відповідає коротка послідовність. На ці короткі послідовності гібридизується флуоресцентні зонди, при цьому кінець одного зонда гасить флуоресценцію флуорофора на початку другого зонда. При цьому для кодування чотирьох основ потрібно всього два типи зондів: кожній основі (а точніше відповідній кожній основі коротка послідовність) відповідає два флуоресцентних сигнали (00, 01, 10, або 11, де 0 відповідає одному кольору, а 1 – іншому). При проходженні через пору



отримана дволанцюгова ДНК розплітається, зонд відділяється, відповідно флуорофор на наступному зонді починає світитись [25].

До переваг метода відноситься точність сигналу – камери реєструють сигнал набагато точніше інших технік. Проте метод потребує попередньої обробки зразка: конвертації кожного нуклеотида в приблизно 12 нуклеотидів (що також подовжує саму ДНК) [25].

### 2.5.3 Порівняння біологічних і твердотільних нанопор

Твердотільні нанопори позбавлені деяких недоліків біологічних нанопор: чутливості до рН, температури, концентрації електроліту, механічних впливів і т.д. Окрім цього, вони стабільні, довше служать, отримати різноманітні форми і розміри таких пор значно легше, а технологія виробництва схожа з виробництвом напівпровідників, що сильно полегшує процес отримання таких пор і робить потенціально можливим поєднання з іншими нанопристроями. До переваг біологічних нанопор можна віднести можливість хімічної або генетичної модифікації, хімічну специфічність до ДНК або РНК і відносно низьку швидкість проходження ДНК або РНК через пору [25, 31, 32].

#### **Інші нанопори**

Для отримання нанопор може бути використана технологія ДНК – оригамі. Вперше така можливість була продемонстрована в 2012 році, коли з допомогою ДНК-оригамі була отримана структура, схожа на альфа – гемолізін. Отримана структура самовільно влаштовувалась в мембрану. Також було показано, що одношарові вуглецеві нанотрубки також можуть влаштовуватись в метрану і пропускати ДНК [25].

На даний момент твердотільні нанопори не володіють хімічною специфічністю білків, тому активно вивчається можливість інтеграції білкових нанопор в твердотільних підкладках. Іншими перспективами напрямку є

використання твердотільних нанопор з сенсорами (ємнісними датчиками, тунельними електронними і іншими детекторами) [25, 31]

## 2.6. Основні протоколи секвенування ДНК

### 2.6.1. Протокол Rapid

Даний протокол використовується для швидкого секвенування в будь-яких умовах. Він дуже швидкий і дозволяє отримувати гігантські прочитання, і є самим портативним протоколом, що потребує мінімум обладнання. Приготування бібліотеки включає два етапи:

- Тагментація ДНК (обробка ДНК транспозонним комплексом, який ріже молекули і одночасно пришиває адаптери.
- Пришивання адаптерів секвенування, які несуть на собі моторні білки.

Для приготування бібліотеки потрібно:

- Певна кількість ДНК
- Два набори реагентів
- Витратні матеріали (пробірки, безнуклеазна вода)
- Центрифуга
- Таймер
- Термоциклер
- Автоматичні піпетки

Щоб отримати довгі прочитання розчини можна перемішувати легким постукуванням по пробірці інакше ДНК рветься і на виході будуть не такі довгі молекули. Також не можна перемішувати гострим накінецьником, лише з широким входом. Для перенесення ДНК з пробірки в пробірку та в секвінатор можна використовувати ці ж наконечники [33].

### 2.6.2. Протокол 1D

Даний протокол відрізняється від попереднього великим виходом, більшою затратою часу на приготування бібліотеки.

В протоколі 1D пропонується поділити ДНК. В цьому протоколі проводиться очистка на магнітних кульках (бідах), фрагменти довше 20 kb погано витягуються і можуть загубитись в ході приготування бібліотеки. Якщо поділити матеріал, приводячи всі молекули приблизно до одного розміру, то втрат буде менше. Найбільш доступний спосіб дробіння – звичайний шприц з голкою. Якщо ми хочемо отримати фрагменти 10-20 kb, то обираємо розмір голки товщиною 21G, то потрібно пропустити розчин ДНК 20 разів [33].

Для обробки кінців потрібні:

- Модуль для відновлення кінців і навішування дезоксиаденозину на 3' кінець – End repair / dA-tailing module
- Мікс для легування адаптерів – Blunt/TA Ligase Mix
- AMPure beads(магнітні кульки) для очистки ДНК
- (FFPE Repair Mix) – суміш для відновлення ДНК з фіксованих зразків, яка виправляє одноланцюгові розриви, окиснені основи, заблоковані 3' кінці і домінуючі цитозини. Таким чином покращує бібліотеку і підвищує довжину прочитань.

Після обробки кінців до нависаючого аденозину легуються адаптери секвенування з моторним білком. Заключним етапом приготування бібліотеки є приєднання зв'язуючого елементу, основною функцією якого є «заманити» молекули ДНК до мембрани, таким чином збільшуючи ймовірність зустрічі молекул і пор [33].

#### **Очистка бібліотеки**

Будь-які ДНКові домішки погіршують вихід секвенування. В протоколі Rapid ми нехтуємо чистотою бібліотеки в користь довгих фрагментів, а протокол 1D дає максимальну кількість даних з не великих фрагментів,

зазвичай довжина становить 10 kb. Тому очистка ДНК від того, що могло залишитись після виділення ДНК і від ферментів, які додаються по ходу протоколу – ключовий момент 1D. Для очистки знадобиться технологія SPRI (Solid Phase Reversible Immobilization) – зв'язування молекул ДНК з магнітними частинками в розчині поліетиленглюколя і хлорид натрія і наступне вивільнення молекул в чистий розчин (рис. 2.6) [33].

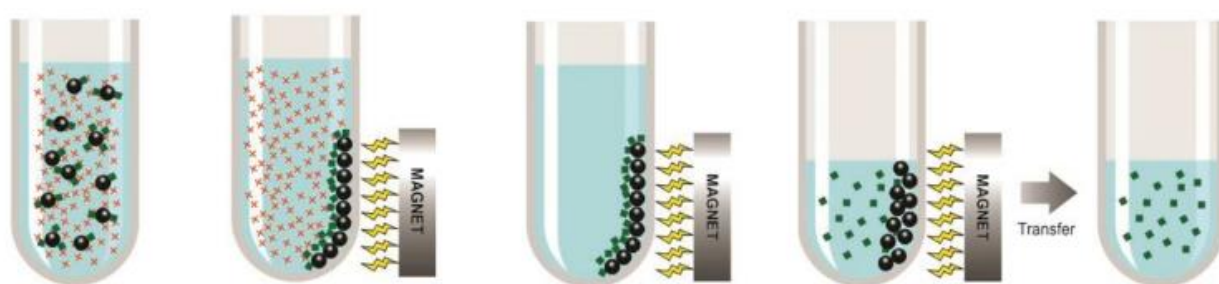


Рисунок 2.6 – Очистка бібліотеки

Біди – чорні магнітні кульки, ДНК – зелені квадратики, непотрібні компоненти – помаранчеві хрестики.

Молекули ДНК сідають на частинки, потім біди концентруються з допомогою магніту і видаляємо розчин з контамінантами (небажані біологічні агенти). Після декількох промивань 70% спиртом молекули ДНК змиваються з бідів, потім біди видаляються з допомогою магніту. Таким чином ми отримуємо бібліотеку вільну від білків, коротких фрагментів ДНК.

Те на скільки короткі фрагменти ДНК виходять можна варіювати з допомогою варіації частинок. Якщо змішуємо бібліотеку з бідами у співвідношенні 1:1,8, відповідно, – позбудемось від фрагментів коротше 100 нуклеотидів, 1:0,4 – коротше 1000 нуклеотидів. Використовувати біди в концентрації менше 0,4 не можна, оскільки занадто низька концентрація хлориду натрія приведе до втрати бібліотеки [33].

В 1D відмивка на бідах виконується декілька разів і ,при розрахунку скільки брати ДНК, потрібно знати, що при кожній відмивці втрачається до

30 % бібліотеки. Також треба враховувати, що молекули довше 20 kb погано змиваються з бідів (електростатична взаємодія), тому довгих прочитань багато не вийде [33].

### 2.6.3 Протокол 1D<sup>2</sup>

Протокол для підвищеної точності секвенування. В попередньому протоколі через пору проходила тільки одна нитка ДНК, компліментарна теж може потрапити в пору і бути відсеквенованою, але це випадковий процес. Протокол 1D<sup>2</sup> модифікований таким чином, що його компліментарна нитка потрапляє в пору слідом за першою і таким чином ми маємо два секвенування одного фрагменту, це сильно зменшує кількість помилкових нуклеотидів (рис.2.7). Заявлена кількість помилок в 1D – 15 %, в 1D<sup>2</sup> – 5 % [33].

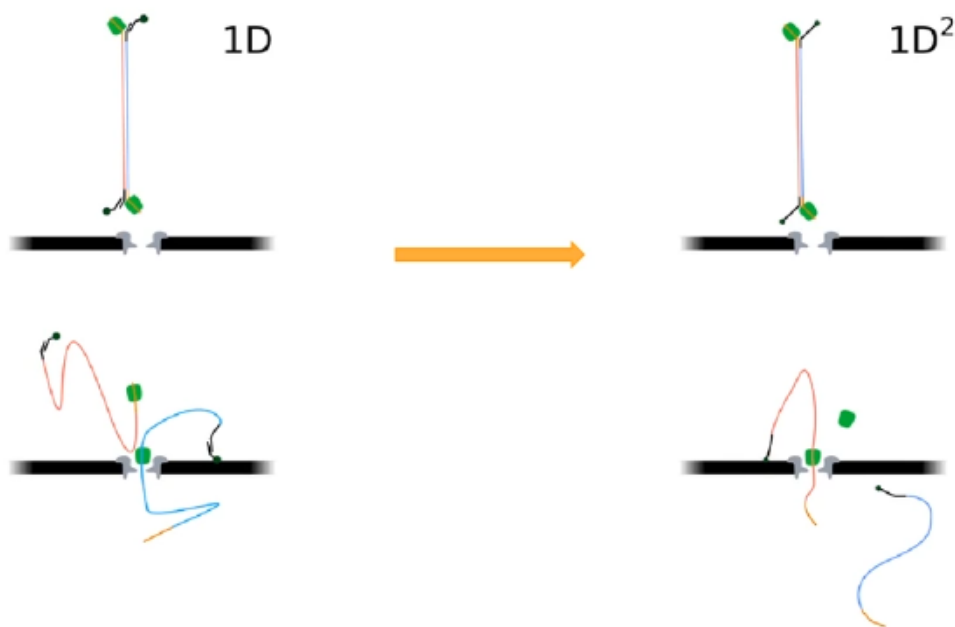


Рисунок 2.7 – Проходження нитки ДНК в протоколах 1D та 1D<sup>2</sup>

Протокол приготування бібліотеки відрізняється від протоколу 1D на один крок – приєднання 1D<sup>2</sup> адаптера [33].

## 2.7. Виділення довгої ДНК

Дуже важливим для секвенування є якість і кількість ДНК.

Вимоги до якості ДНК:

- Чистота ДНК по Nanor (прилад для визначення концентрацій і співвідношення поглинання на довгих хвилях):
  - OD 260/280 в межах 1,8 – 2,0 , якість ДНК хороша
  - OD 260/230 не нижче 2,0
- Без детергентів і інших ПАВ (поверхневі активні речовини)
- Розмір фрагментів > 30 kb (ДНК високомолекулярне)
- Співвідношення концентрацій згідно Qubit (прилад для визначення концентрацій) і Nanodrop повинно бути не більше, ніж 1 : 1,5

Як визначити чистоту по показам Nanodrop:

- OD 260/280 < 1,8 – ДНК забруднена білками, фенолом
- OD 260/280 > 2,0 – є домішка РНК
- OD 260/230 < 2,0 – розчин містить домішки солями, або фенолом, або вуглеводами і краще таку ДНК не пускати в хід.

Чому чистота ДНК так важлива? Білки і інші домішки закупорюють пори і роблять їх не доступними для секвенування.

Довжина ДНК, скільки ДНК потрібно для протоколів без ПЦР? Довжина прочитання приблизно дорівнює довжині фрагмента ДНК.

Для протокола:

- Rapid потрібно – 400 нг ДНК
- 1D – 1-2 мкг
- 1D<sup>2</sup> – 1-2 мкг

Але потрібно брати в 3-4 рази більше ДНК для гарного виходу секвенування, ніж вказано в протоколах.

Методи для виділення довгої ДНК:

- За допомогою колонок, близько 30 kb

- Genomic tips, 50 – 100 kb
- Фенол – хлороформна екстракція, 200 – 300 kb

## 2.8. Яким чином нанопорове секвенування допоможе при вивченні раку

Портативний секвенатор, дозволяючий отримати довгі прочитання в реальному часі – використовується для секвенування повного геному людини в різноманітних галузях людини [34].

Онкологічні дослідження з пристроєм MinION (минайн):

- Цільове ресеквенування з високим покриттям для аналізу гермінальних або соматичних варіантів з використанням технологій збагачення методами гібридизації або мультиплексної ПЛР
- Аналіз локусів, асоційованих з раком
- Використання готових панелей
- Секвенування РНК для аналізу варіантів сплайсингу і химерних транскриптів
- Секвенування цілого геному з використанням довгих рідів дозволяє порівнювати пухлинні і нормальні геноми
- Вивчення структурних варіантів
- Виявлення раніше невідомих однонуклеотидних варіацій SNV (Single Nucleotide Variation) [34].

## 2.9 Висновки до розділу 2

Нанопорове секвенування має ряд переваг: висока точність довгих прочитань, дешевизна, висока чутливість, мобільність, швидкий аналіз і відображення результатів в реальному часі. За допомогою секвенування можна виявити будь – яку зміну в генетичному коді людини, що є вірним шляхом до доступного раннього діагностування різних хвороб на етапі зміни послідовності в ДНК. В розділі описано процес секвенування через нанопору, наведено приклади біологічних та твердотільних пор, розглянуто принципово важливі аспекти для секвенування такі як: кількість і якість ДНК, описано основні протоколи, що використовують в секвенуванні нового покоління.



### 3 МЕТОДИ ВИМІРЮВАННЯ МЕМБРАННОГО ПОТЕНЦІАЛУ

#### 3.1 Метод фіксації потенціалу

Для вивчення потенціалзалежних мембранних каналів застосовують метод фіксації потенціалу. В такому методі використовують електронну систему зі зворотнім зв'язком. Така система забезпечує автоматичне підтримання мембранного потенціалу. Різницю потенціалів фіксують на певному рівні по різні сторони мембрани, при цьому мембранний потенціал можна змінювати на суворо певну величину. Таким методом можна вимірювати іонні струми, що протікають крізь мембрану через канали, котрі активуються при зміні потенціалу. Якщо зміна напруги на мембрані стала, то відповідно до закону Ома зміна струму однозначно пов'язана зі зміною провідності [35].

Технічна фіксація потенціалу здійснюється наступним чином. Використовуючи підсилювач – регулятор внутрішньоклітинний потенціал порівнюють з керуючим потенціалом. Будь – яке відхилення мембранного потенціалу від керуючого підсилюється і на виході підсилювача виникає керуючий струм. Такий струм протікає через електроди, що розташовані по різні сторони мембрани в такому напрямку, що мембранний потенціал знову стає рівний керуючому. Це автоматичне узгодження стається за долю мілісекунди після того, як задається керуючий потенціал. Коли у відповідь на таку ступінчасту деполяризацію відкривається канал, відповідні іони входять в пору по електрохімічному градієнту і переносять з собою електричні заряди. Заряди, що входять прагнуть зсунути мембранний потенціал в позитивному напрямку, але навіть мінімальне відхилення від керуючого потенціалу одразу компенсується в результаті видалення з клітин надлишкових зарядів з допомогою підсилювача – регулятора. При цьому реєструється певний струм, який подається підсилювачем для підтримання мембранного потенціалу на необхідному рівні, цей струм в точності рівний іонному струму, що протікає через мембрану [35].

### 3.2 Петч – кламп метод

Для вимірювання іонних струмів через одиничні канали спочатку було запропоновано непрямий метод для аналізу мембранного шуму. Згодом був розроблений спосіб прямої реєстрації одиничних іонних каналів з допомогою методу, що називається петч – кламп. В сукупності ці підходи дають відповіді на питання, що стосуються функції іонних каналів: який заряд проходить через одиночний іонний канал, наскільки довго по часу канал залишається відкритим, як час знаходження іонного каналу в закритому чи відкритому стані залежить від мембранного потенціалу [36].

Петч – клемп – один з методів електрофізіології, дозволяючий ізолювати фрагмент клітинної мембрани з іонними каналами, задавати певну різницю потенціалів через цей фрагмент клітинної мембрани, створювати по обидві сторони мембрани середовище з певним складом і вимірювати, при цих гарно контрольованих умовах, струм іонних каналів. Метод дозволяє спостерігати поведінку і хімічне перетворення одиничних молекул. Різні модифікації методу дозволяють в експерименті варіювати змінними, що впливають на поведінку каналів [37].

Петч – кламп метод значно розширює область знань про функціонування іонних каналів. Для петч – кламп реєстрування потрібно, щоб кінчик скляної піпетки зі згладженими краями з внутрішнім діаметром близько 1 мкм щільно контактував з мембраною клітини, що досліджують. При вдалому підводі, завдяки легкому присмоктуванню, між клітинною мембраною і склом піпетки буде створюватися опір більше  $10^9$  Ом. Фіксуються досить невеликі струми, що проходять через ділянку мембрани, яка знаходиться всередині кінчика піпетки в момент, коли піпетка з'єднана з певним підсилювачем. Наведена конфігурація петч – кламп метода називається cell attached (контакт з клітиною). Високоомний контакт гарантує, що іонні струми, що проводяться цим фрагментом клітинної мембрани не втрачаються в місці контакту петч –

піпетки з клітиною, а проходять переважно через підсилюючу апаратуру. При використанні петч – кламп метода реєстровані події складаються з струмових сигналів у вигляді прямокутних сигналів, що відображають процеси відкриття і закриття одиночних іонних каналів. Таким чином, в реальному часі можемо спостерігати активність білкових молекул мембрани[36].

Коли вивчають зміну провідності однотипних каналів у відповідь на якісь хімічні впливи чи залежність їх провідності від різниці потенціалів на мембрані, зручно фіксувати потенціал і виміряти струм. Це найчастіший варіант петч – кламп, називаються *voltage clamp*. Але, інколи дослідника цікавлять процеси трансмембранного переносу іонів, пов'язаних зі зміною мембранного потенціалу. В таких випадках фіксують на постійному рівні струм і вивчають зміну різниці потенціалів, даний варіант називається *current clamp*.

В простому випадку струми одиночних каналів з'являються нерегулярно і мають різну тривалість, але з постійною амплітудою (рис. 3.1 А). Деякі іонні канали, візьмемо приклад каналу в відкритому стані, можуть мати більш, ніж один рівень провідності, приклад показано на (рис. 3.1 В). Окрім цього, іонні канали можуть проявляти комплексну кінетику, наприклад, струм через одиночний іонний канал може виглядати не як простий прямокутник, а як «спалах» відкритий канал (рис. 3.1 С) [36].

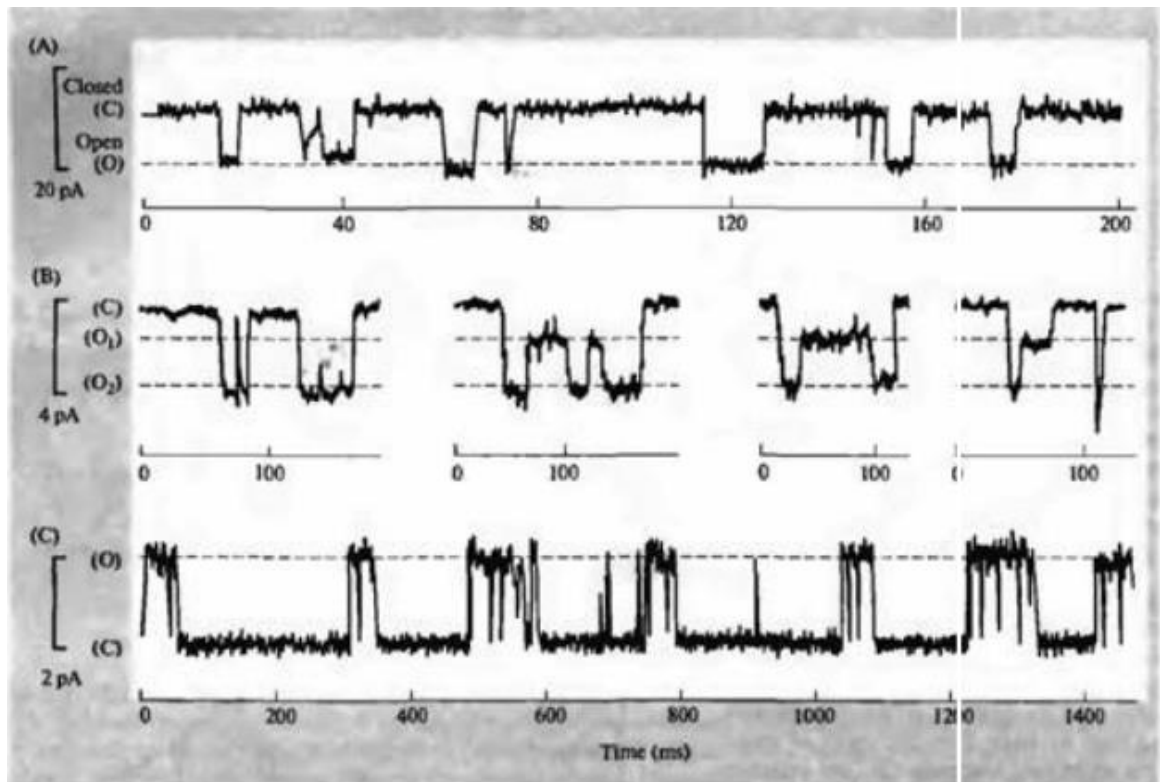


Рисунок 3.1 – Приклади іонних каналів, зареєстрованих методом петч – кламп

Таким чином, петч – кламп метод надає нові можливості для вивчення поведінки іонних каналів. По – перше, ізоляція маленької ділянки мембрани дозволяє спостерігати активність всього декількох іонних каналів, а не тисячі, які активуються в усій клітині. По – друге, високий опір контакту надає можливість реєструвати навіть дуже малі струми. В результаті ми маємо можливість точного виміру амплітуд струмів одиночних іонних каналів і можемо провести аналіз кінетики каналів [36].

Метод дозволяє:

1. Проводити багато досліджень в межах класичної електрофізіології
2. Реєструвати струми і потенціали від клітин дуже малих розмірів (3 – 10 мкм)
3. Реєструвати струми одиночних каналів амплітудою порядку пікоампер (пА)
4. Досліджувати дію лікувальних препаратів при швидкому підводі їх до сторони мембрани, як до зовнішньої, так і до внутрішньої [38].

### 3.3 Висновки до розділу 3

У цьому розділі розглянуто основні методи для реєстрації зміни струму на мембрані. Зміна струму чи потенціалу на мембрані здійснюється методом петч –кламп. В даній темі розглядається дослідження з фіксацією напруги.

#### 4 РОЗРОБКА АЛГОРИТМІВ ДЛЯ МОДЕЛЮВАННЯ ТА АНАЛІЗУ СИГНАЛІВ НАНОПОРОВОГО СЕКВЕНУВАННЯ ДНК

Аналіз біологічних послідовностей вимагає попередньої обробки сигналів нанопорового секвенування ДНК з метою усунення шуму та знаходження однакових сегментів у послідовностях, а також ділянок з мутаціями. Крім того знаходження однакових послідовностей є важливим для відстеження співпадань у базах даних. Важливою задачею для розуміння мутацій є вирівнювання однакових біологічних послідовностей та виявлення ділянок з мутаціями та визначення типів цих мутацій.

Вирівнювання послідовностей ДНК – це розташування первинних послідовностей ДНК або амінокислот різних зразків з метою виявлення областей подібності між ними. Це допомагає виявити зв'язки між послідовностями. Таким чином, вирівнювання послідовностей ДНК зазвичай виконується для того, щоб визначити функції та структуру нових секвенованих генів шляхом порівняння з геном, функція якого відома. Результатом такого порівняння може бути визначення еволюційного зв'язку між послідовностями внаслідок мутацій або локально однакових областей у віддалених послідовностях. Таке порівняння може призвести до виявлення точної відповідності або близької відповідності між послідовністю секвенованих нуклеотидів і послідовностями, які вже досліджені і зберігаються в базах даних. Точна відповідність ділянок може спостерігатися по всій довжині послідовності нуклеотидів (глобальна відповідність) або для менших сегментів послідовності нуклеотидів (локальна відповідність).

З метою розробки алгоритму знаходження мутацій у ДНК тканин, уражених раком, в даній роботі використовувалися фрагменти послідовностей нуклеотидів з наступних публічних баз даних:

1) COSMIC – catalogue of somatic mutation in cancer (база даних соматичних мутацій знайдених при раку)[39];



Норма:

GGAGGGGAGAACTTGCCACTTTGGCTGAGTTGGTCCAGTATTAC

Нонсенс:

GGAGGGTAGAACTTGCCACTTTGGCTGAGTTGGTCCAGTATTAC

Місенс:

GGAGGGGAGAAATTTGCCACTTTGGCTGAGTTGGTCCAGTATTAC

Делеція:

GGAGAGAACTTGCCACTTTGGCTGAGTTGGTCCAGTATTACGGG

Інсерція

GGATGCGGGGAGAACTTGCCACTTTGGCTGAGTTGGTCCAGTAT

З метою застосування методів аналізу сигналів до задачі знаходження мутацій дані послідовності нуклеотидів було використано для моделювання сигналів, отриманих за допомогою нанопорового секвенування. Проходження кожного нуклеотиду крізь білкову нанопору змінює значення рівня іонного струму. Негативно заряджена одноланцюгова ДНК проходить крізь нанопору в мембрані, зовнішня поверхня якої має негативний заряд, а внутрішня – позитивний. Як тільки черговий нуклеотид перекриває внутрішній отвір в нанопорі, електропровідність мембрани (струм мембрани) змінюється в залежності від хімічної будови нуклеотиду. Нуклеотиди різного типу, які складають ланцюжок ДНК, відрізняються хімічними та фізичними властивостями. Тому ступінь блокування нанопори кожним нуклеотидом різна. Відповідно до цього, змінюються електрична провідність та струм мембрани. Відомо, що під час проходження ДНК крізь нанопору іонний струм значно зменшується.

При моделюванні сигналу струму крізь нанопору під час протягування одноланцюгової ДНК використовувалися наступні параметри, які відповідають реальним патч-кламп установкам у режимі фіксації напруги:

- Частота дискретизації сигналу 5 кГц;
- Швидкість протягування ДНК крізь нанопору – 500 нуклеотидів/с;



- Середній струм іонів крізь нанопору у випадку перекриття пори гуаніном складає 10пА, у випадку перекриття тиміном - 20пА, у випадку перекриття аденіном - 30пА, у випадку перекриття цитозином - 40пА;
- Тривалість знаходження нуклеотиду в порі складає 2 мс. Загалом сигнал, при проходженні крізь пору 45 нуклеотидів, за умови  $F_s = 5$  кГц, має тривалість 90 мс та кількість відліків 450;
- Час дискретизації - 0,2 мс (кожний нуклеотид моделюється як сходинка, яка включає в себе десять відліків сигналу).

Знаходження певного нуклеотиду в нанопорі моделюється як сходинка тривалістю  $\tau = 2$  мс, що символізує час блокування нанопори 1 нуклеотидом (рис. 4.2).

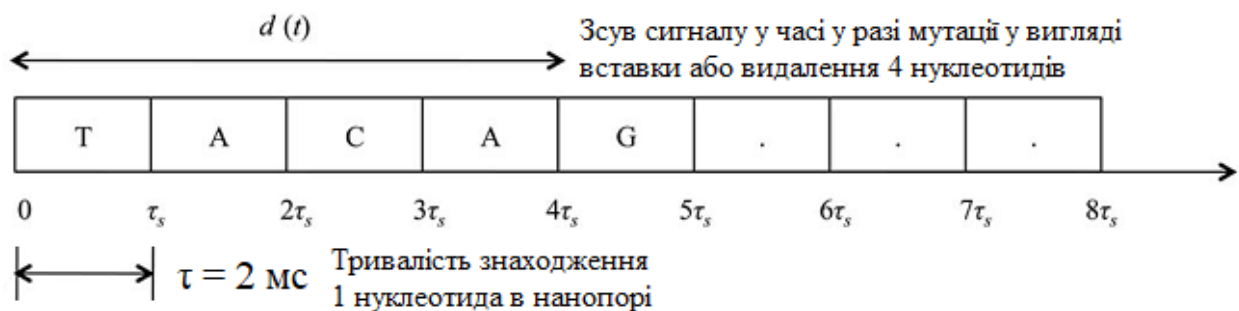


Рисунок 4.2 – Пояснення до моделювання сигналу струму крізь нанопору під час протягування одноланцюгової ДНК з мутаціями інсерції або делеції

Серед параметрів моделі струму струм крізь нанопору у разі проходження одного нуклеотиду слід відзначити наступні:

- $d_{st}$  – час початку секвенування крізь нанопору;
- $\tau$  – тривалість знаходження 1 нуклеотида в порі;
- $c_{st} = d_{st} + \tau/2$  – центр сходинки струму іонів крізь нанопору;
- $I_{free}$  – струм іонів крізь вільну нанопору (без ДНК ланцюга);
- $A$  – амплітуда прямокутного імпульсу, яка відповідає за зменшення струму крізь нанопору за умови знаходження в ній певного нуклеотиду;

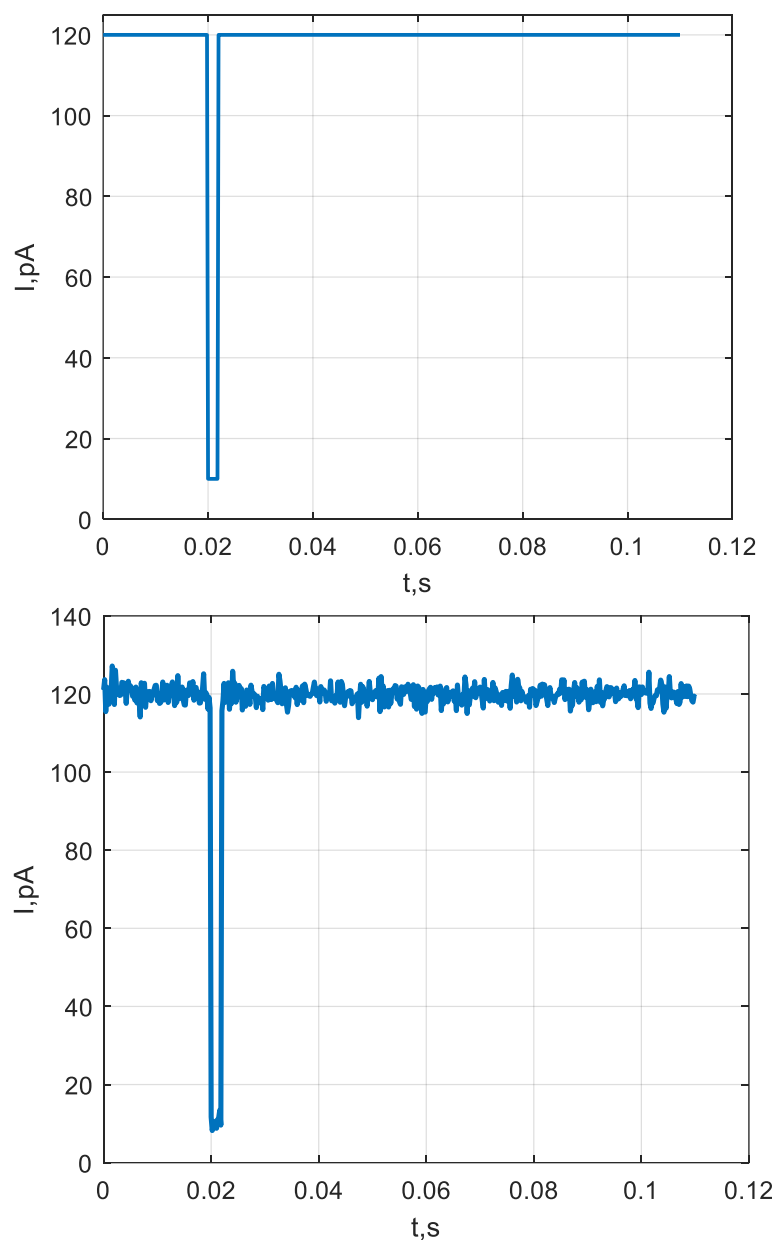


Рисунок 4.3 – Модельний сигнал струму крізь нанопору при блокуванні нанопори одним нуклеотидом гуаніна без шуму та з наявністю білого шуму відповідно

З метою отримання реалістичного сигналу струму крізь нанопору під час протягування одноланцюгової ДНК необхідно додати змодельований білий шум до отриманого сигналу (рис. 4.3 б).

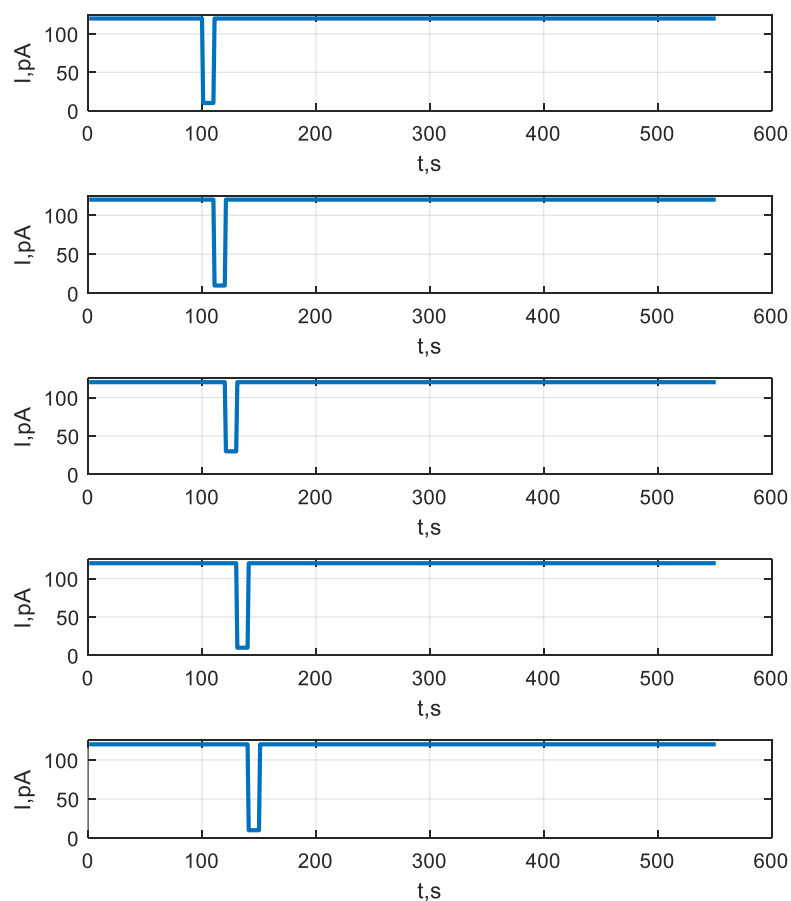
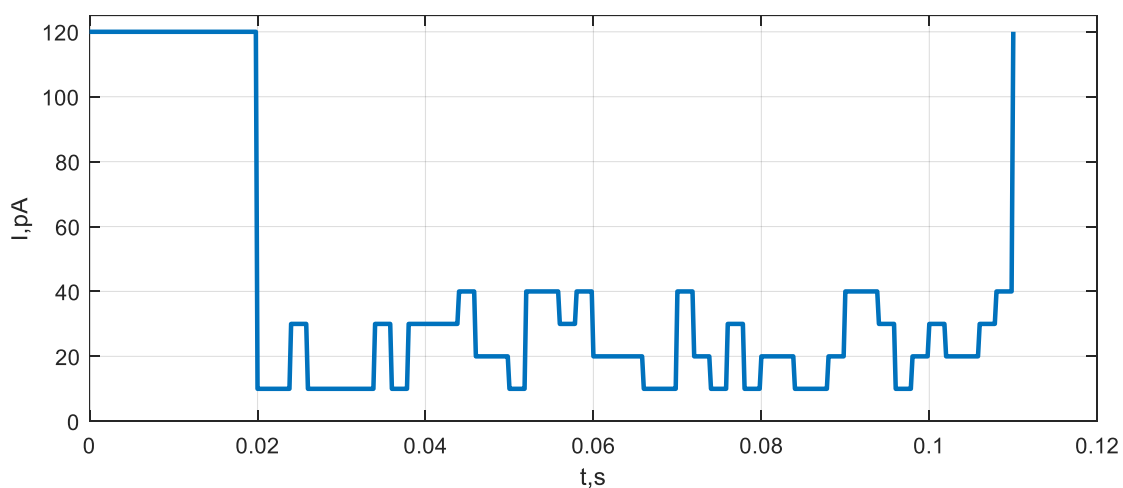
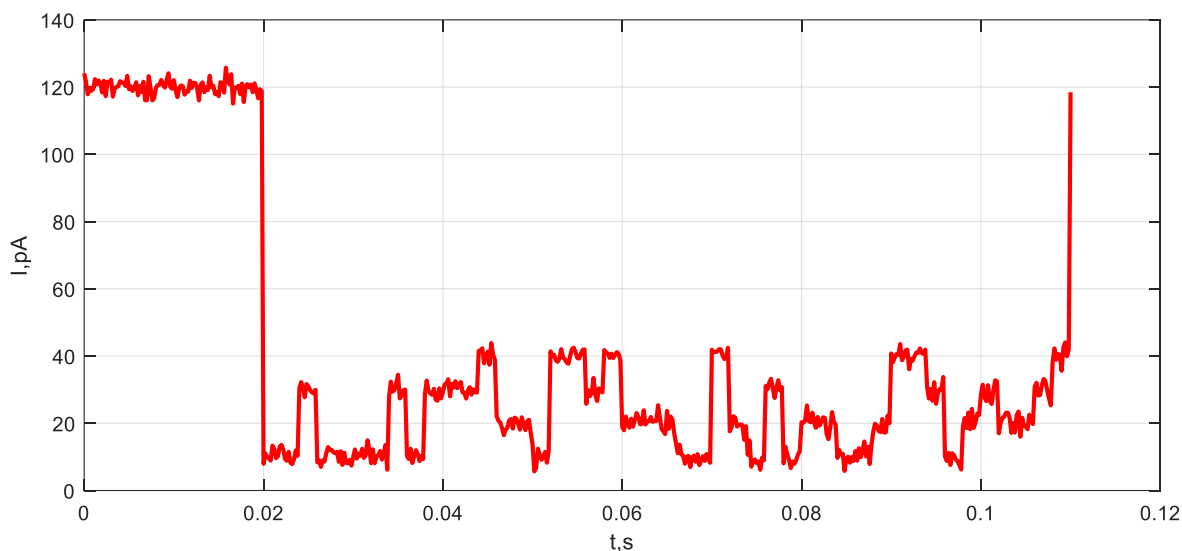


Рисунок 4.4 – Приклад моделювання сходинок струму крізь нанопору при послідовному блокуванні нанопори 5-ти нуклеотидами GGAGG



GGAGGGGAGAACTTGCCACTTTGGCTGAGTTGGTCCAGTATTAC (варіант «норма» – відсутність мутацій).

Рисунок 4.5 – Приклад моделювання струму крізь нанопору при послідовному блокуванні нанопори 45-ти нуклеотидами



GGAGGGGAGAACTTGCCACTTTGGCTGAGTTGGTCCAGTATTAC (варіант “норма – відсутність мутацій).

Рисунок 4.6 – Приклад моделювання струму крізь нанопору при послідовному блокуванні нанопори 45-ти нуклеотидами з наявністю білого шуму

Таким чином, було змодельовано низку сигналів в нормі та при наявності ряду мутацій, серед яких на рисунках 4.7 – 4.10 наведено:

- точкова нонсенс – мутація, варіант з шумом і без шуму (рис. 4.7 – 4.8).  
Стрілочкою вказано місце мутації – однонуклеотидну заміну.
- делеція – мутація з видаленням одного чи ряду нуклеотидів (рис. 4.9 – 4.10).

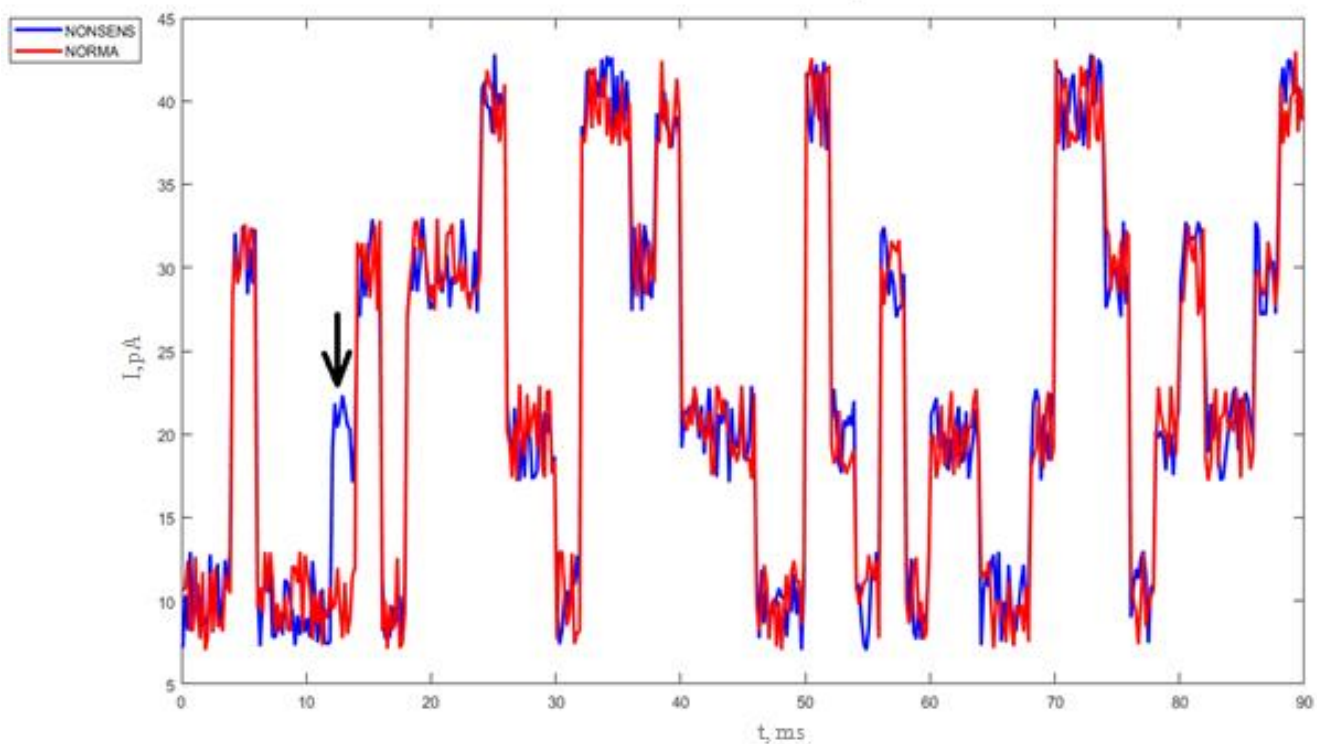


Рисунок 4.7 – Нонсенс – мутація, сигнал з шумом і початковий модельний сигнал в нормі

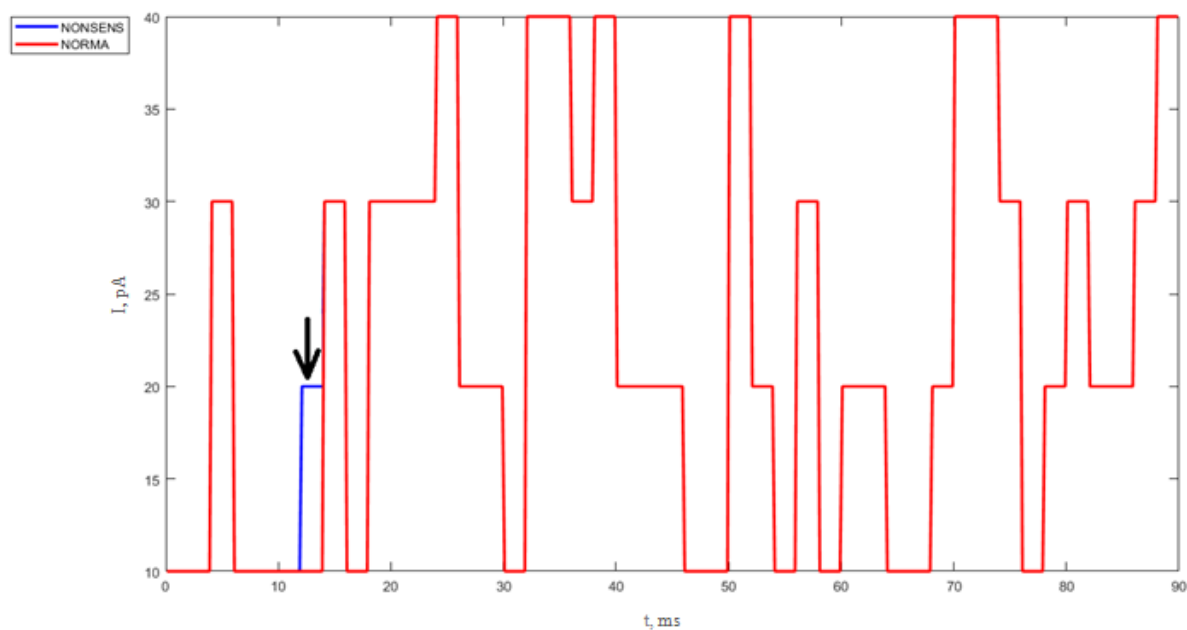


Рисунок 4.8 – Нонсенс – мутація, чистий сигнал і початковий модельний сигнал в нормі

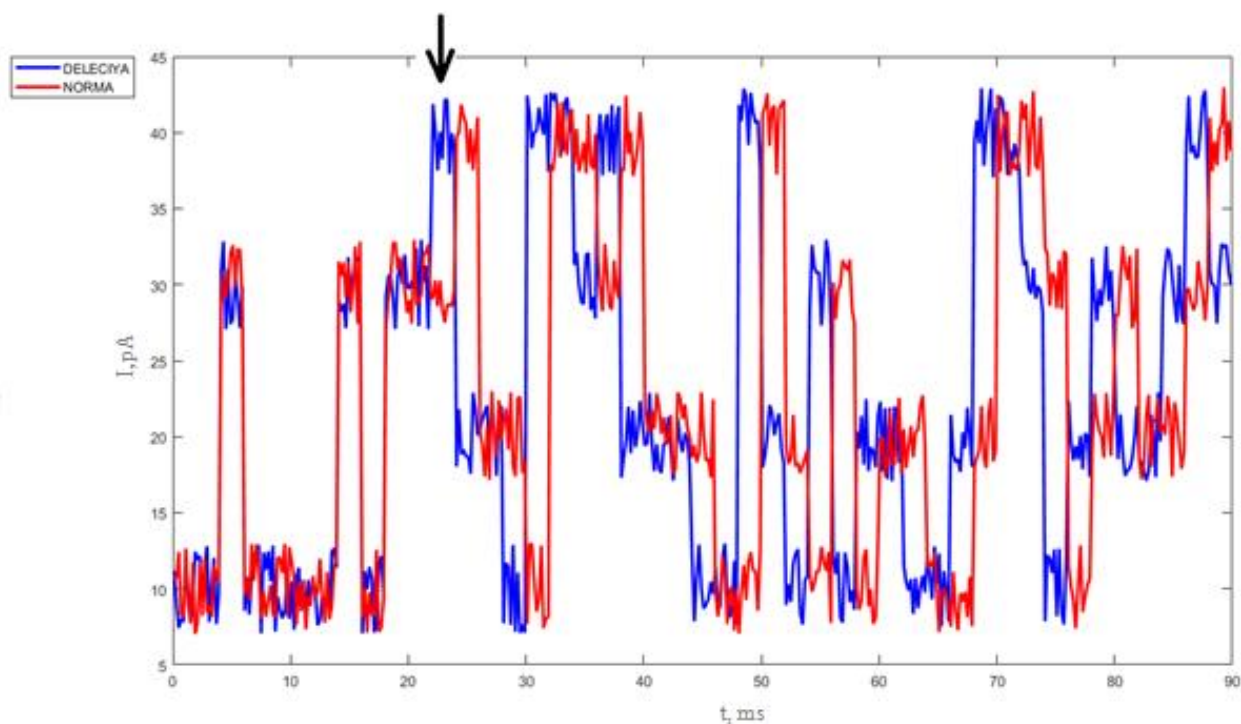


Рисунок 4.9 – Делеція – видалення ділянки ДНК, як наслідок - зсув послідовності вліво, сигнал з шумом і початковий модельний сигнал в нормі

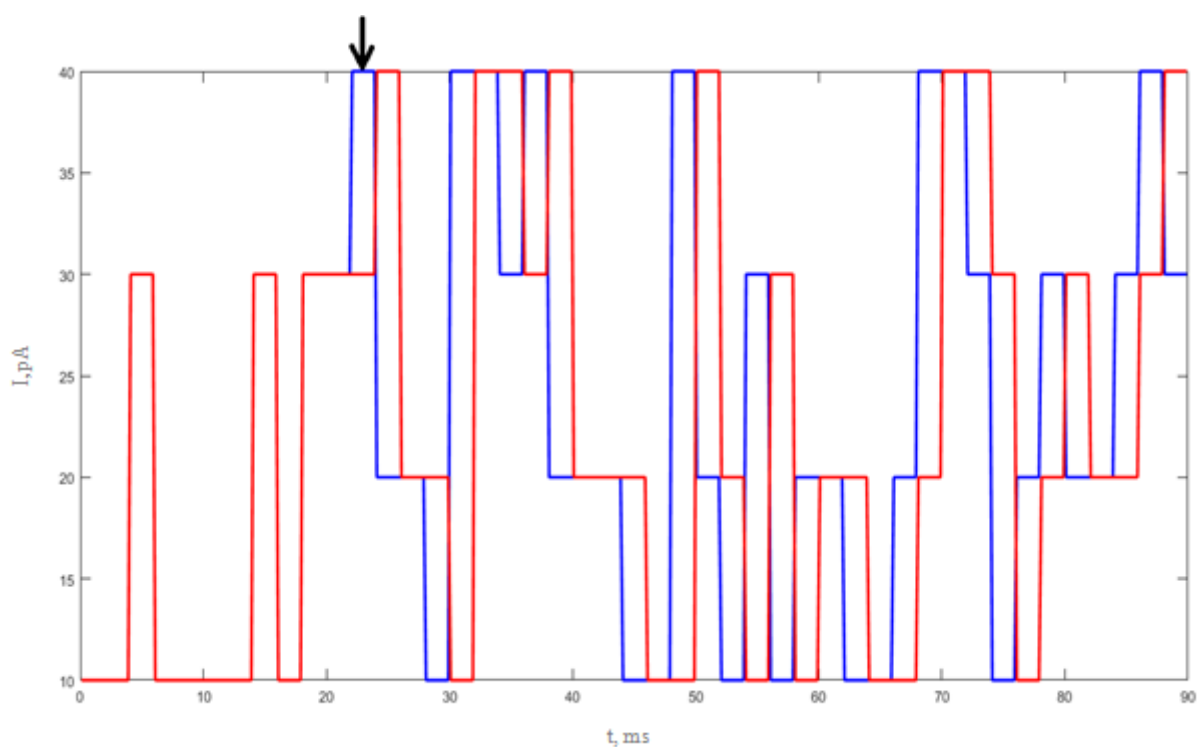


Рисунок 4.10 – Делеція – видалення ділянки ДНК, як наслідок - зсув послідовності вліво, чистий сигнал і початковий модельний сигнал в нормі

#### 4.1 Пошук подібних областей між двома і більше послідовностями методом кореляційного аналізу

Кореляційний аналіз грає велику роль в теорії сигналів. Сенс цього аналізу в кількісному вимірюванні ступеню схожості різних сигналів. В цьому допомагає кореляційна функція, що представляє собою інтеграл від добутку двох копій сигналу, що зсунуті один відносно одного на час  $\tau$ , у випадку автокореляції, наприклад, пошук схожих фрагментів (послідовності нуклеотидів) в одному й тому ж самому сигналі [42].

$$B_S(\tau) = \int_{-\infty}^{\infty} s(t)s(t - \tau)dt. \quad (4.1)$$

Кореляційна функція (КФ) показує ступінь схожості між сигналом і зсунутою копією цього сигналу. Тобто, більше значення кореляційної функції – сильніша схожість [42].

Якщо КФ показує ступінь схожості між зсунутими копіями одного і того ж сигналу, то взаємна кореляційна функція (ВКФ; англійський термін – cross – correlation function, CCF) дозволяє виміряти аналогічну величину для зсунутих екземплярів двох різних сигналів. В даній роботі ми порівнюємо два сигнали двох нуклеотидних послідовностей. Загальний вид формули КФ зберігається, але під інтегралом добуток двох різних сигналів, один з яких з затримкою на час  $\tau$  [42]:

$$B_{12}(\tau) = \int_{-\infty}^{\infty} s_1(t)s_2(t - \tau)dt. \quad (4.2)$$

Область вирівнювання послідовностей нуклеотидів полягає у пошуку областей подібності між двома або більше послідовностями. В даній роботі пропонується застосування кореляційного аналізу для визначення подібності за

допомогою кореляційної функції між двома формами сигналів струму у нанопорі (в нормі та з наявністю мутації).

Необхідно знайти пік кореляційної функції та його зсув відносно початку координат (у відліках та за часом). По значенню зсуву, тобто позитивний він чи від'ємний, можна визначити в яку сторону потрібно зсунути сигнал і який це тип мутації (інерція чи делеція). Для інсерції буде зсув вправо, оскільки відбувається вставка нуклеотидів, для делеції вліво – видалення нуклеотидів. Відстань між відліками – 0,2 мс. Тоді для інсерції, де вставка 3 нуклеотиди, тобто зсув у 30 відліків – 0,006 с маємо (рис. 4.11):

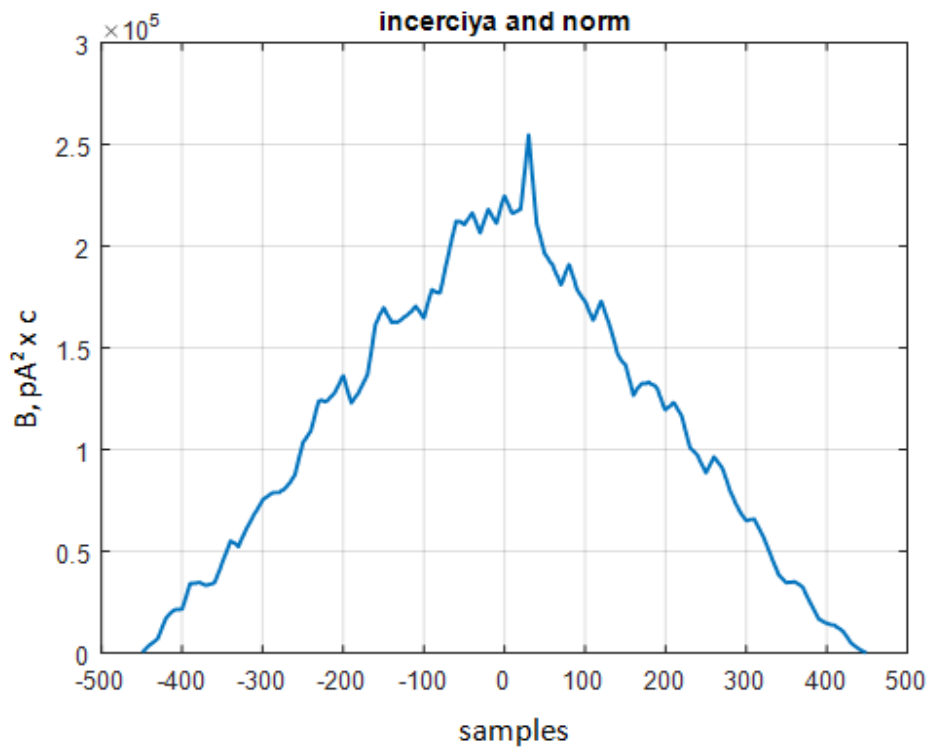


Рисунок 4.11 – Пік кореляційної функції між сигналом в нормі та сигналом з мутацією інсерції 3 зайвих нуклеотидів - зсув 30 відліків вправо від початку координат



Теж саме для делеції, варіант з видаленням одного і трьох нуклеотидів (рис. 4.12 – 4.13):

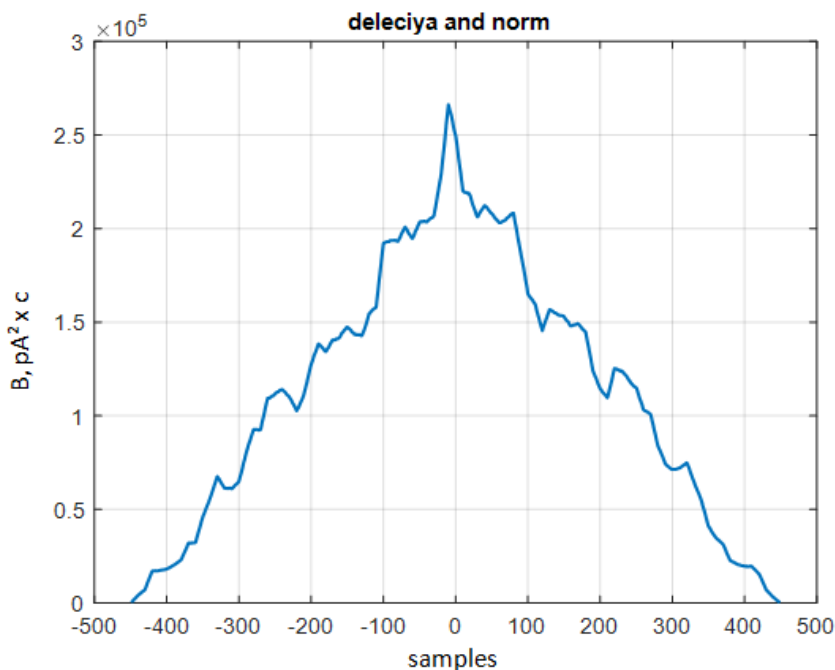


Рисунок 4.12 – Пік кореляційної функції між сигналом в нормі та сигналом з мутацією делеції 1го нуклеотида - зсув 10 відліків вліво від початку координат

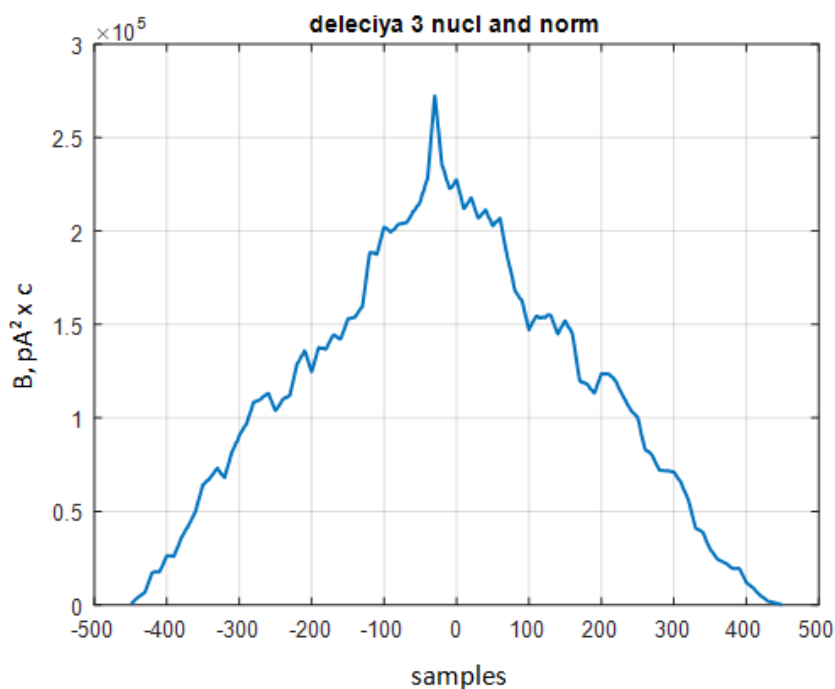


Рисунок 4.13 – Пік кореляційної функції між сигналом в нормі та сигналом з мутацією делеції 3х нуклеотидів - зсув 30 відліків вліво від початку координат

Нижче наведено результати після накладання співпадаючих фрагментів для всіх трьох варіантів (рис.4.14 – 4.16):

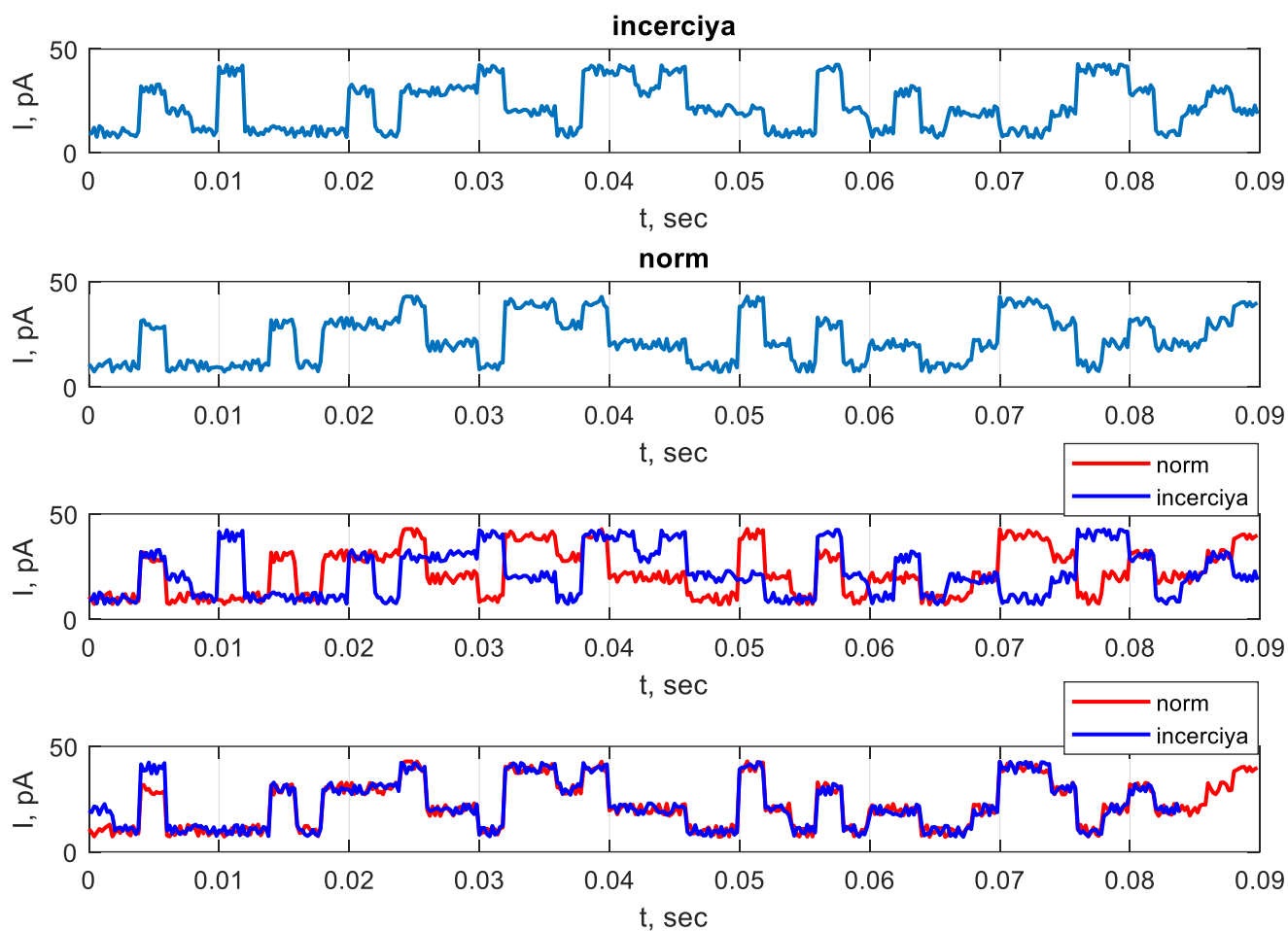


Рисунок 4.14 – Накладання співпадаючих фрагментів сигналу в нормі та сигналу з мутацією інсерції 1го нуклеотида

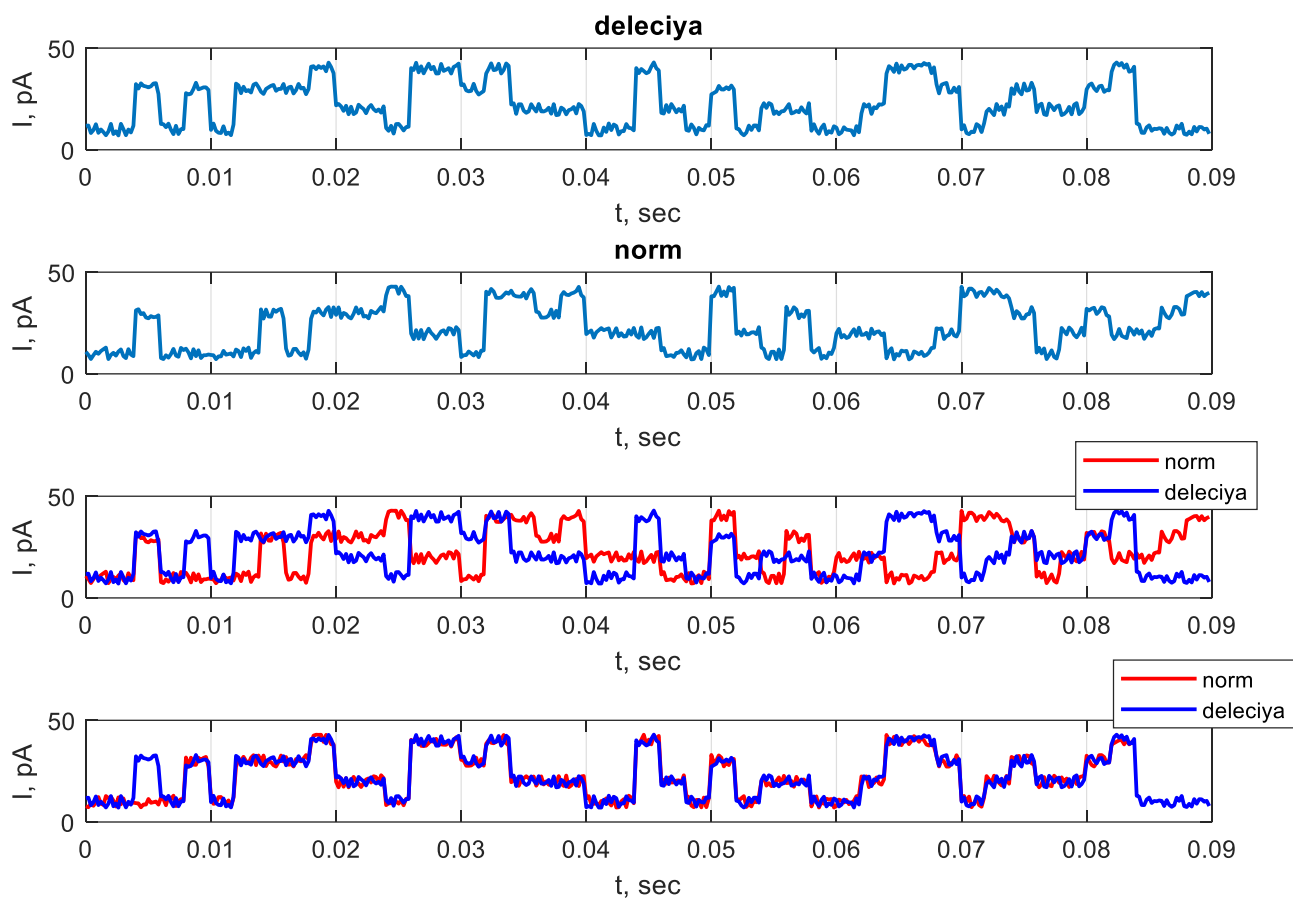


Рисунок 4.15 - Накладання співпадаючих фрагментів сигналу в нормі та сигналу з мутацією делеції 1го нуклеотида

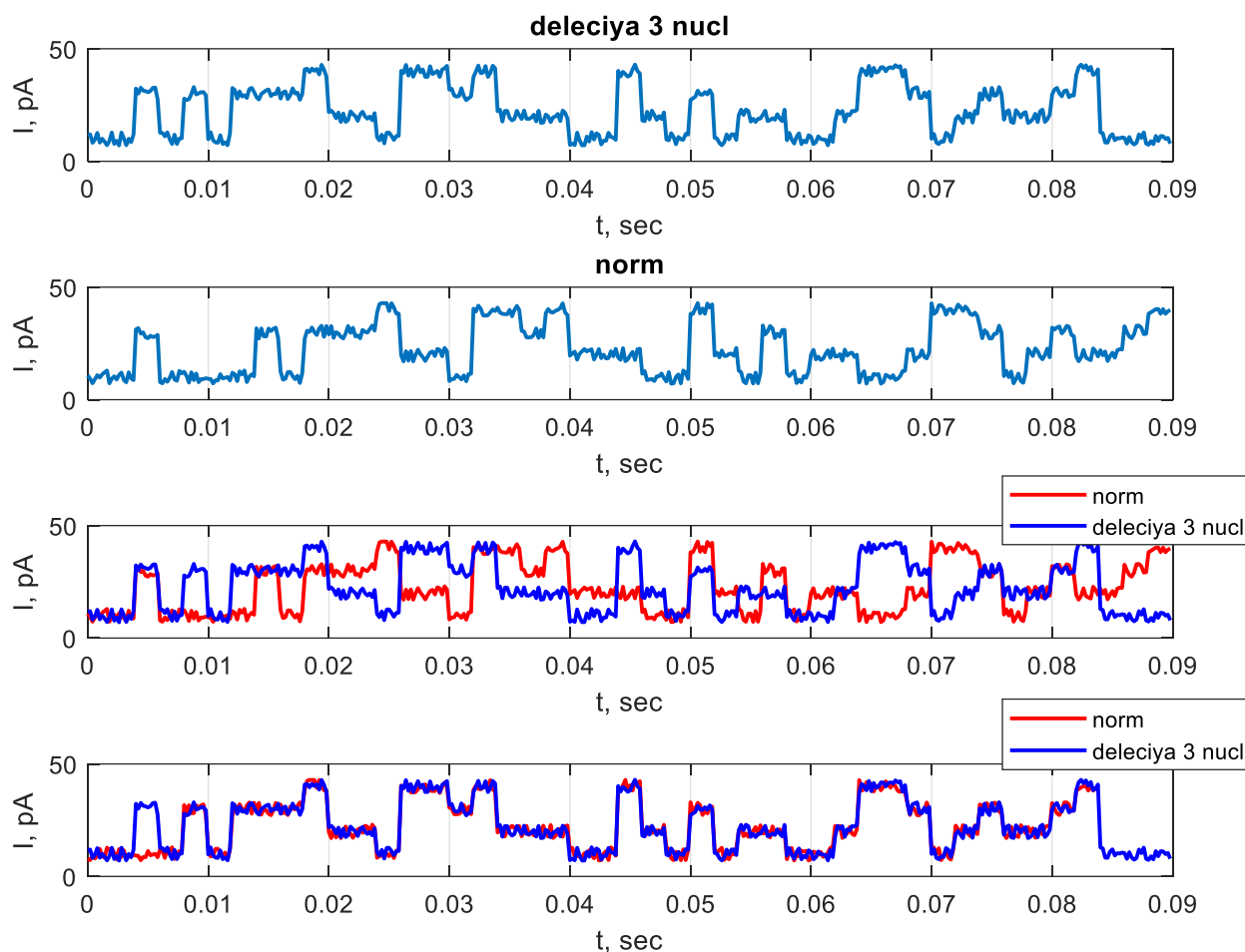


Рисунок 4.16 – Накладання співпадаючих фрагментів сигналу в нормі та сигналу з мутацією делеції 3х нуклеотидів

#### 4.2 Алгоритм знешумлення сигналів.

В реальності, сигнали з установки патч-кламп реєструються з великими шумами, для подальшого розпізнавання нуклеотидів сигнал потрібно очистити від шумів. У разі частотної фільтрації з застосуванням фільтру нижніх частот або у разі застосування вейвлет-розкладу для зменшення рівня шуму, отримуємо сигнал із спотворенням сходинок, які передають струм у нанопорі під час блокування її нуклеотидами. Це пояснюється тим, що спектр прямокутного сигналу має частотні складові на всьому діапазоні частот від 0 до

$F_s/2$ . Видалення із сигналу складових високих частот викликає спотворення фронтів прямокутних сходинок, що в свою чергу призводить до некоректного розшифрування послідовності нуклеотидів (рис. 4.17 – 4.18).

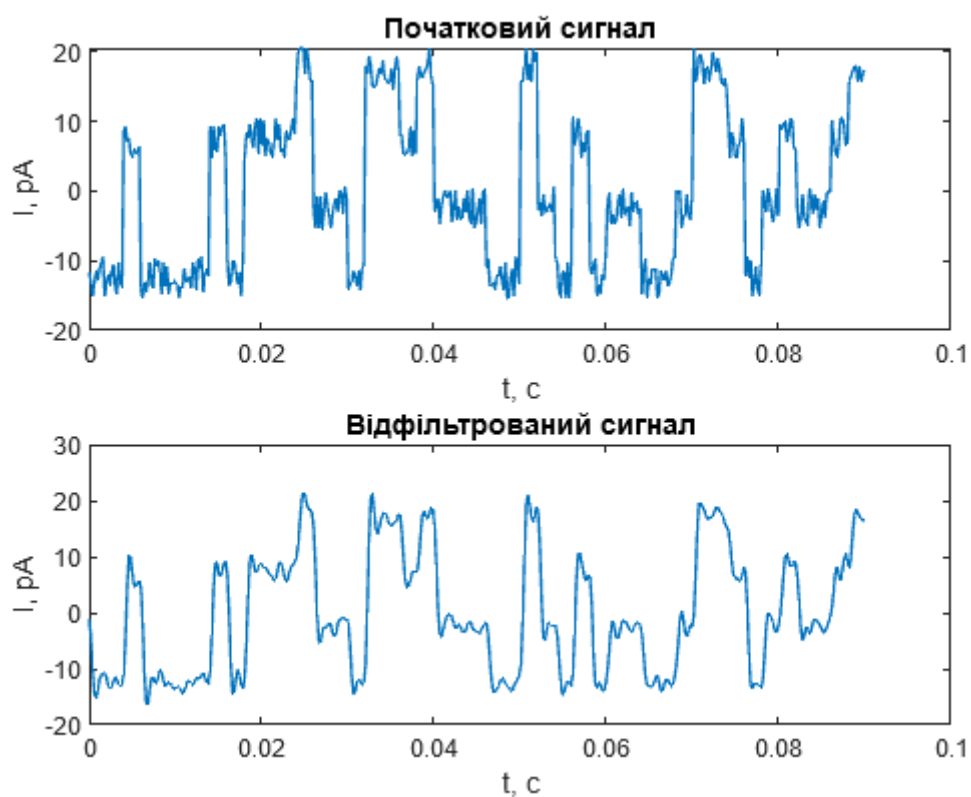


Рисунок 4.17 – Зменшення рівня шуму фільтром Баттерворта

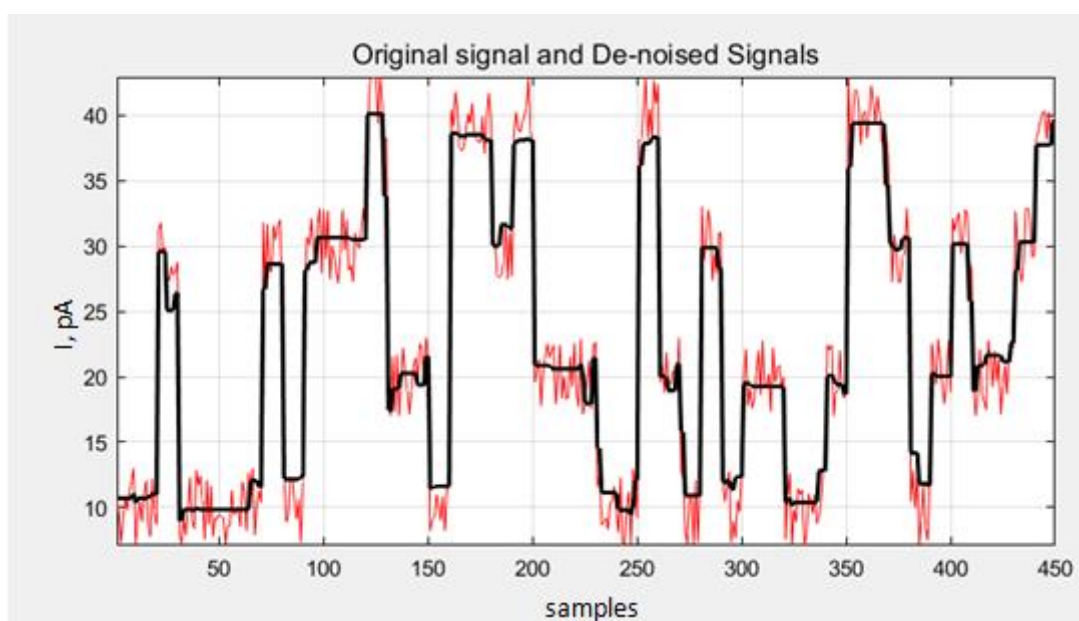


Рисунок 4.18 – Зменшення рівня шуму за допомогою порогової обробки коефіцієнтів вейвлет-розкладу

Тому у даній роботі для учення шумових складових пропонується наступний алгоритм:

1. Фіксуємо початок входження чергового нуклеотида в нанопору. Створюємо часове вікно довжиною в тривалість знаходження 1 нуклеотида в нанопорі ( $\tau = 2$  мс, що за умови  $F_s=5$  кГц відповідає 10 відлікам сигналу).
2. Знаходимо середнє значення струму крізь нанопору під час знаходження в ній поточного нуклеотиду. Так як білий шум має нульове середнє значення, то середнє значення струму крізь нанопору під час знаходження в ній поточного нуклеотиду буде наближатися до фіксованих значень, властивих саме цьому нуклеотиду.
3. Знаходимо модуль різниці між вектором значень струму крізь нанопору, властивим для 4 нуклеотидів А, Т, G, С та отриманим середнім значенням на часовому інтервалі вікна. Далі знаходимо, де ця різниця є мінімальною і присвоюємо всій ділянці струму у часовому вікні знайдене фіксоване значення.
4. Процедуру повторюємо для всього сигналу, зсуваючи часове вікно вздовж сигналу на крок, що відповідає тривалісті знаходження 1 нуклеотида в нанопорі ( $\tau = 2$  мс, що за умови  $F_s=5$  кГц відповідає 10 відлікам сигналу).
5. Отримуємо сходинковий сигнал без шуму (рис. 4.18).
6. Згідно зі значеннями сходинок струму повертаємо код з літер, які відповідають послідовності нуклеотидів секвенованої ділянки одноланцюгової ДНК.

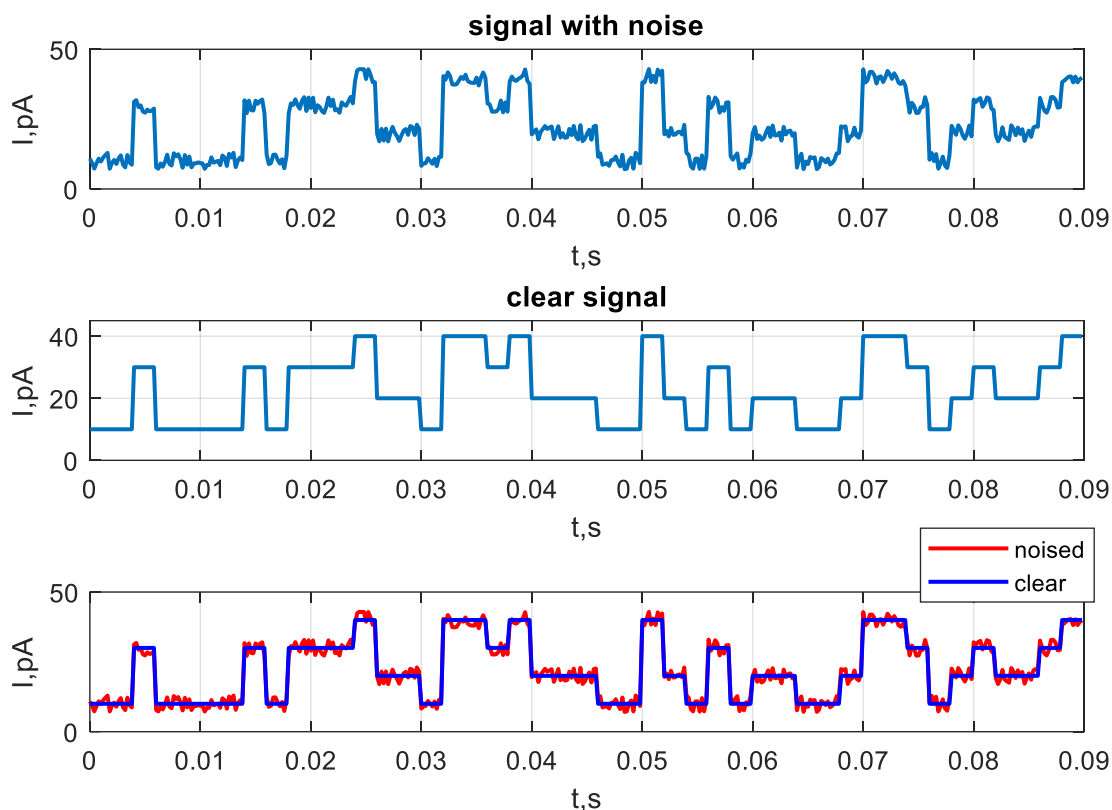


Рисунок 4.19 – Очищення сигналів від шумів вище описаним алгоритмом

### 4.3 Класифікація сигналів нанопорового секвенування

З метою дослідження можливості застосування алгоритмів машинного навчання до задачі класифікації сигналів нанопорового секвенування було змодельовано по 300 сигналів відрізка послідовності нуклеотидів для 4х класів:

- 1) норма,
- 2) мутація у вигляді зміни 1 нуклеотида (5й нуклеотид А змінено на Т),
- 3) делеція 3х нуклеотидів,
- 4) інерція 3х нуклеотидів.

До змодельованого сигналу кожного класу додавався адитивний гаусівський шум з нульовим середнім.

Алгоритм  $k$  найближчих сусідів, є одним з найпростіших алгоритмів машинного навчання. Побудова моделі полягає в запам'ятовуванні навчального набору даних. Для того, щоб зробити прогноз для нової точки даних, алгоритм знаходить найближчі до неї точки навчального набору, тобто знаходить «найближчих сусідів».

У найпростішому варіанті алгоритм  $k$  найближчих сусідів розглядає лише одного найближчого сусіда - точку навчального набору, найближче розташовану до точки, для якої потрібно отримати прогноз. Але можна розглянути будь-яку кількість ( $k$ ) сусідів. Коли приймають до уваги більше одного сусіда, для присвоєння мітки використовується клас, який найбільш часто зустрічається: іншими словами, обирається клас, який набрав більшість серед  $k$  найближчих сусідів. У разі мультикласової класифікації підраховується кількість сусідів, що належать до кожного класу, і прогнозується клас, який найбільш часто зустрічається.

В класифікаторі **KNeighbors** є два важливі параметри: кількість сусідів і міра відстані між точками даних. У даній роботі використовувався параметр кількість сусідів = 10 та евклідова відстань, яка між сигналами  $x$  і  $y$  в  $n$ -вимірному просторі обчислюється за формулою (4.3):

$$d(x,y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad .(4.3)$$

Результати класифікації суттєво залежать від рівня шуму та типу мутації. При відносно низькому рівні шуму, коли значення сходинок струму для різних нуклеотидів не перекриваються, точність класифікації досягає 100%. Але у разі збільшення амплітуди змодельованих шумових складових, виникає перекриття класів для норми і одонуклеотидної мутації, особливо якщо рівні сходинок для 2х нуклеотидів в нормі і при мутації лежать поруч (як у випадку А та Т). У цьому випадку виникають помилки класифікації через перекриття рівнів сигналу для 1 нуклеотиду. У разі мутацій делеції або існерції декількох нуклеотидів буде зсув між сигналом в нормі та при мутації. Через цей зсув буде



зростати евклідова відстань між сигналами. Тому ці класи класифікатор виявляє без помилок.

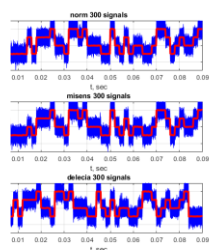


Рисунок 4.20 – Змодельовані сигнали нуклеотидних послідовностей (по 300 реалізацій для кожного випадку: норма, місенс мутація, інсерція, делеція)

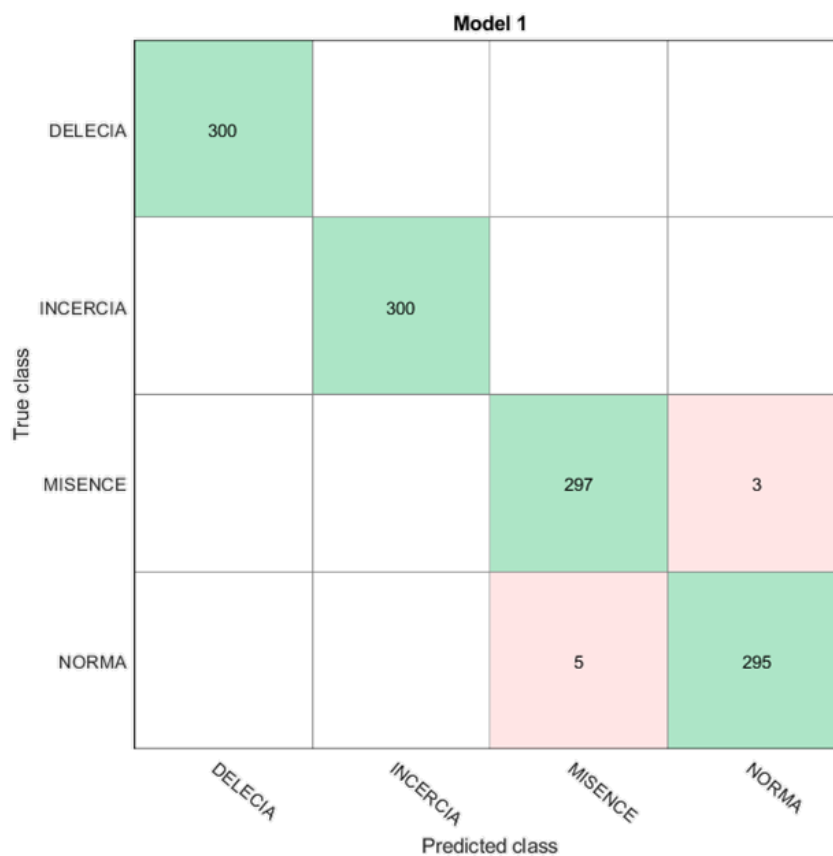


Рисунок 4.21 – Матриця похибок класифікації

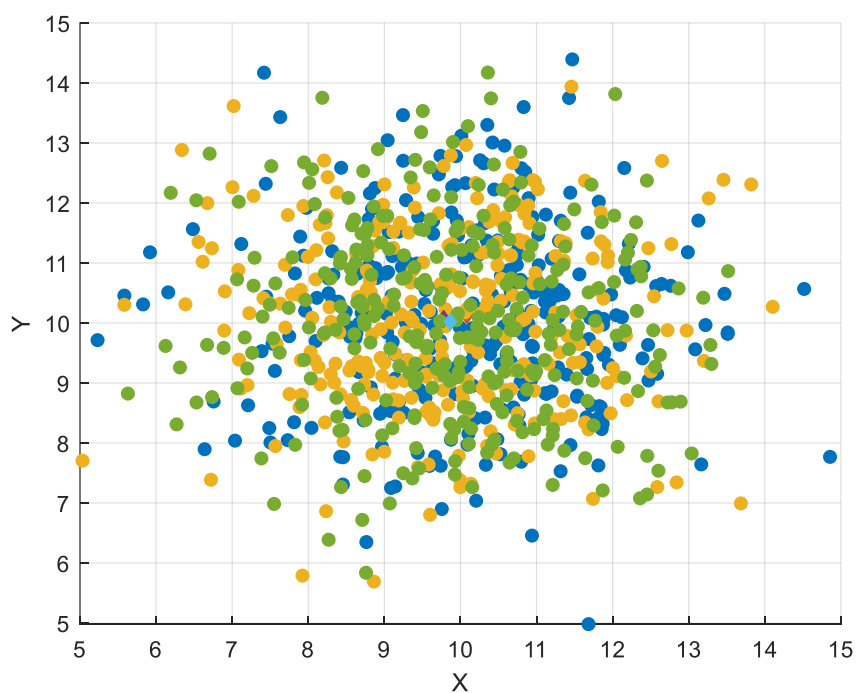


Рисунок 4.22 – Приклад розкиду значень сигналу нуклеотидної послідовності для 3 класів "норма", "делеція", "інсерція" в момент часу, коли співпав вид нуклеотиду (10 pA)

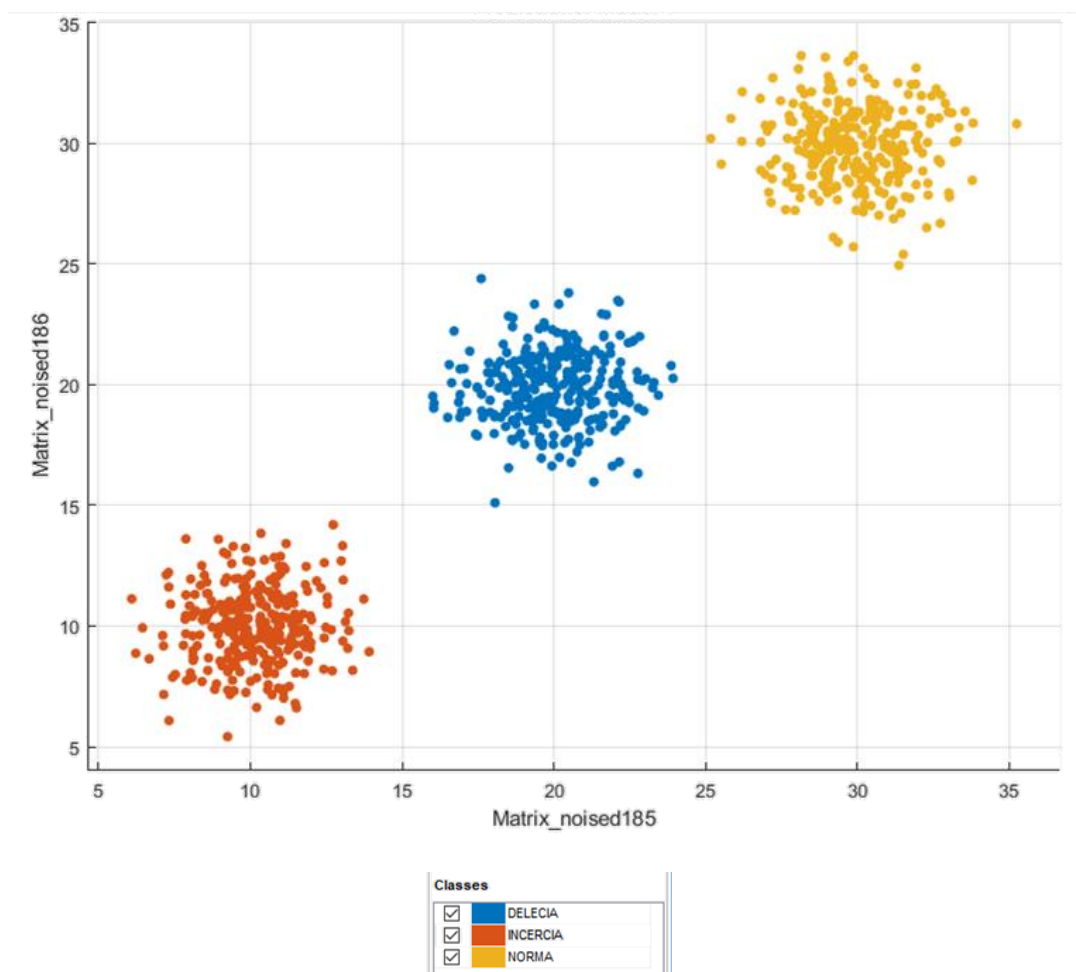


Рисунок 4.22 - Приклад розкиду значень сигналу нуклеотидної послідовності для 3 класів "норма", "делеція", "інсерція" в момент часу, коли не співпали види нуклеотидів (10, 20, 30 pA)

#### 4.4 Висновки до розділу 4

Проведено роботу по обробці попередньо змодельованого сигналу з метою усунення шумів, знаходження однакових сегментів у послідовностях, а також ділянок з мутаціями.

Наведено приклади порівняння сигналів при нормі і патології, продемонстровано зміну іонного струму на ділянці, коли пора відкрита і в момент проходження через неї нуклеотидів.

Для знаходження однакових фрагментів між двома сигналами застосовували кореляційний аналіз. Для знешумлення сигналу скористались методом вейвлет – розкладу та алгоритмом, за яким знаходиться середнє значення струму кожного нуклеотиду, таке середнє значення порівнюється з заданими значеннями векторів, обирається найближче значення (що і буде дорівнювати величині іонного струму одного з нуклеотидів) і фрагмент замінюється на чисту сходинку.

Результати класифікації показують, що рівень шуму суттєво впливає на точність класифікації. При досягненні певного рівня шуму відбувається перекриття класів. Такі мутації як інсерції і делеції класифікуються без помилок, тому що через зсув між сигналом в нормі та при мутації буде зростати евклідова відстань між цими сигналами. Тому ці класи класифікатор виявляє без помилок.

## ВИСНОВКИ

Було розроблено алгоритми для моделювання та аналізу сигналів нанопорового секвенування. Наведено приклади мутацій в порівняльній формі норма – патологія, де вказано зміну в послідовності. Застосовано кореляційний аналіз для визначення подібності за допомогою кореляційної функції між двома формами сигналів струму у нанопорі. По піку кореляції можна визначити, який це тип мутації (інсерція, делеція) та на скільки відліків зсунувся сигнал. Для подальшого розпізнавання нуклеотидів сигнал був очищений від шумів. У разі частотної фільтрації з застосуванням фільтру нижніх частот або у разі застосування вейвлет-розкладу (рис. 4.17) для зменшення рівня шуму, отримуємо сигнал із спотворенням сходинок, тому було обрано інший алгоритм, детально описаний у розділі 4.2. В роботі використано класифікатор **KNeighbors**. Результати класифікації суттєво залежать від рівня шуму та типу мутації. При відносно низькому рівні шуму, коли значення сходинок струму для різних нуклеотидів не перекриваються, точність класифікації досягає 100%. У разі збільшення амплітуди змодельованих шумових складових, виникає перекриття класів для норми і одонуклеотидної мутації, особливо якщо рівні сходинок для 2х нуклеотидів в нормі і при мутації лежать поруч (як у випадку А та Т). У цьому випадку виникають помилки класифікації через перекриття рівнів сигналу. Інсерції і делеції класифікуються без помилок, оскільки цим мутаціям характерний зсув в декілька нуклеотидів між сигналами в нормі і патології, через який зростає евклідова відстань.

## ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Науколандия. Строение ДНК. – Режим доступу до ресурсу : <https://scienceland.info/biology10/dna-structure>
2. Beyondcurriculum. Почему природа выбрала именно фосфаты. – Режим доступу до ресурсу : <https://blog.beyondcurriculum.kz/why-nature-chose-phosphates/>
3. Химимк. Вторичная структура нуклеиновых кислот. – Режим доступу до ресурсу : <https://www.xumuk.ru/biologhim/039.html>
4. Новокшанова, А. Л. Биохимия для технологов : в 2 ч. / А. Л. Новокшанова. – М. : Юрайт, 2018. Ч. 1 – 211 с. – ISBN 978-5-534-10322-9.
5. Яккласс. Генетический код. – Режим доступу до ресурсу : [https://www.yaklass.ru/p/biologia/obschie-biologicheskie-zakonomernosti/biokhimicheskie-protCESSy-v-kletke-16037/khranenie-i-peredacha-geneticheskoi-informatcii-dnk-i-rnk-geneticheskii-k\\_-17333/re-56bb55b9-c072-481d-9ef6-3579ea21d0e6](https://www.yaklass.ru/p/biologia/obschie-biologicheskie-zakonomernosti/biokhimicheskie-protCESSy-v-kletke-16037/khranenie-i-peredacha-geneticheskoi-informatcii-dnk-i-rnk-geneticheskii-k_-17333/re-56bb55b9-c072-481d-9ef6-3579ea21d0e6)
6. Український медичний часопис. Уровень лейцина влияет на рост опухолей. – Режим доступу до ресурсу : <https://www.umj.com.ua/article/141638/uroven-lejtsina-vliyaet-na-rost-opuholej>
7. Геномед. Соматические мутации в опухоли. – Режим доступу до ресурсу : <https://genomed.ru/somaticheskie-mutacii-v-opuholi/>
8. Newtonew. Генетическое тестирование в онкологии. – Режим доступу до ресурсу : <https://newtonew.com/science/preduprezhden-i-vooruzhen-geneticheskoe-testirovanie-v-onkologii>
9. Рідлі, М. Геном. Автобіографія у 23 главах / М. Рідлі. – К. : Видавнича група КМ – БУКС, 2018. – 408 с. – ISBN 978-617-7489-67-1.

10. Postnauka. Изменения генома при раке. – Режим доступа до ресурсу :  
<https://postnauka.ru/faq/12933>
11. Элементы. На пути к детальному каталогу раковых генов. – Режим  
 доступа до ресурсу :  
[https://elementy.ru/novosti\\_nauki/432446/Na\\_puti\\_k\\_detalnomu\\_katalogu\\_rakovykh\\_genov](https://elementy.ru/novosti_nauki/432446/Na_puti_k_detalnomu_katalogu_rakovykh_genov)
12. Nature biotechnology. Differences that matter in cancer genomics. –  
 Режим доступа до ресурсу : <https://www.nature.com/articles/nbt.2715>
13. Cell. Mapping the Hallmarks of Lung Adenocarcinoma with Massively  
 Paralleled Sequencing. – Режим доступа до ресурсу :  
<https://www.cell.com/fulltext/S0092-8674%2812%2901061-6>
14. Nature. Discovery and saturation analysis of cancer genes across 21 tumour  
 types. – Режим доступа до ресурсу :  
<https://www.nature.com/articles/nature12912>
15. Элементы. Мутации в ассоциированных с раком генах не обязательно  
 приводят к раку. – Режим доступа до ресурсу :  
[https://elementy.ru/novosti\\_nauki/433489/Mutatsii\\_v\\_assotsiirovannykh\\_s\\_rakom\\_genakh\\_ne\\_obyazatelno\\_privodyat\\_k\\_raku](https://elementy.ru/novosti_nauki/433489/Mutatsii_v_assotsiirovannykh_s_rakom_genakh_ne_obyazatelno_privodyat_k_raku)
16. Элементы. При раке молочной железы основные мутации  
 накапливаются уже в предраковом состоянии. – Режим доступа до  
 ресурсу :  
[https://elementy.ru/novosti\\_nauki/433248/Pri\\_rake\\_molochnoy\\_zhelezy\\_osnovnye\\_mutatsii\\_nakaplivayutsya\\_eshche\\_v\\_predrakovom\\_sostoyanii/t5271886/Onkologiya](https://elementy.ru/novosti_nauki/433248/Pri_rake_molochnoy_zhelezy_osnovnye_mutatsii_nakaplivayutsya_eshche_v_predrakovom_sostoyanii/t5271886/Onkologiya)
17. Indicator. Секвенирование ДНК. – Режим доступа до ресурсу :  
<https://indicator.ru/label/sekvenirovanie-dnk>
18. Oxford Nanopore Technologies. – Режим доступа :  
<https://nanoporetech.com/>
19. Diam. Нанопоровое секвенирование Oxford Nanopore. Режим доступа  
 до ресурсу : <https://www.dia-m.ru/news.php?newsid=59771>

20. SkyGen. Уникальная технология нанопорового секвенирования Oxford Nanopore Technologies. – Режим доступа до ресурсу : [https://www.skygen.com/podderzhka/obzory/42\\_unikalnaya\\_tehnologiya\\_nanoporovogo\\_sekvenirovaniya\\_oxford\\_nanopore\\_technologies/](https://www.skygen.com/podderzhka/obzory/42_unikalnaya_tehnologiya_nanoporovogo_sekvenirovaniya_oxford_nanopore_technologies/)
21. Пропионикс. Общие сведения о секвенировании билполимеров. – Режим доступа до ресурсу : <http://propionix.ru/sekvenirovaniye-i-ptsr#ngs>
22. Lenta. Последовательность ДНК определили с помощью нанопор. – Режим доступа до ресурсу : <https://lenta.ru/news/2012/09/24/newnanoseq/>
23. Stepik. Сравнение возможностей секвенирования 2-го и 3-го поколения. Режим доступа до ресурсу : <https://stepik.org/lesson/83316/step/5?discussion=1546933&unit=59947>
24. Вечная молодость. Секвенирование ДНК. Нанопора с моторчиком. – Режим доступа до ресурсу : <http://vechnayamolodost.ru/articles/nanotekhnologii/sekdnnasmo1e/http://vechnayamolodost.ru/articles/nanotekhnologii/sekdnnasmo1e/>
25. PNAS. Single – nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. – Режим доступа до ресурсу : <https://www.pnas.org/content/106/19/7702>
26. INFOX. Разработан метод рас шифровки ДНК без секвенирования. – Режим доступа до ресурсу : <https://www.infox.ru/news/174/56235-razrabotan-metod-rassifrovki-dnk-bez-sekvenirovania>
27. Элементы. Немецкие биотехнологии создали из ДНК искусственные ионные каналы. – Режим доступа до ресурсу : [https://elementy.ru/novosti\\_nauki/431938/Nemetskie\\_biotekhnologi\\_sozdali\\_iz\\_DNK\\_iskusstvennye\\_ionnye\\_kanalny](https://elementy.ru/novosti_nauki/431938/Nemetskie_biotekhnologi_sozdali_iz_DNK_iskusstvennye_ionnye_kanalny)
28. PNAS. Nanopore DNA sequencing with MspA. – Режим доступа до ресурсу : <https://www.pnas.org/content/107/37/16060>



29. FIND-INFO. Биологические моторы. – Режим доступа до ресурсу :  
<http://find-info.ru/doc/dictionary/nanotechnology/articles/199/biologicheskie-motory.htm>
30. Вечная молодость. Новый метод секвенирования ДНК. – Режим доступа до ресурсу :  
<http://vechnayamolodost.ru/articles/nanotekhnologii/sekdnnasmole/http://vechnayamolodost.ru/articles/nanotekhnologii/sekdnnasmole/>
31. Oxford Nanopore Technologies. Types of nanopores. Режим доступа до ресурсу : <https://nanoporetech.com/how-it-works/types-of-nanopores>
32. Hindawi. Solid – State Nanopore Based DNA Sequencing Technology. – Режим доступа до ресурсу :  
<https://www.hindawi.com/journals/jnm/2016/5284786/>
33. Stepic. Основные протоколы секвенирования ДНК. – Режим доступа до ресурсу : <https://stepik.org/lesson/83320/step/9?unit=59951>
34. SkyGen. Области применения нанопорового секвенатора MinION. – Режим доступа до ресурсу :  
[https://www.skygen.com/podderzhka/obzory/43\\_oblasti\\_primeneniya\\_nanoporovogo\\_sekvenatora\\_minion/](https://www.skygen.com/podderzhka/obzory/43_oblasti_primeneniya_nanoporovogo_sekvenatora_minion/)
35. Vseobiology. Метод измерения биопотенциалов. – Режим доступа до ресурсу : <https://vseobiology.ru/biofizika/1235-041-metody-izmereniya-biopotentsialov-metody-fiksatsii-napryazheniya-petch-klyamp-metod-vnutrikletochnoj-perfuzii>
36. Николлс, Д. От нейрона к мозгу / Д. Николлс, Р. Мартин, Б. Валлас, П. Фукс. – М. : Едиториал УРСС, 2003. – 672 с. – ISBN 5-354-00162-5.
37. Livejournal. Patch – Clamp. – Режим доступа до ресурсу :  
<https://hoegni.livejournal.com/33516.html>
38. Научная Библиотека им. А. Н. Игнатова. Методы измерения ионных токов. – Режим доступа до ресурсу : [http://www.f-mx.ru/biologiya/metody\\_izmereniya\\_ionnyx\\_tokov.html](http://www.f-mx.ru/biologiya/metody_izmereniya_ionnyx_tokov.html)

39. Catalogue Of Somatic Mutations In Cancer. – Режим доступа до ресурсу : <https://cancer.sanger.ac.uk/cosmic>
40. National Center for Biotechnology Information. – Режим доступний до ресурсу : <https://www.ncbi.nlm.nih.gov/>
41. National Cancer Institute. – Режим доступа до ресурсу : <https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>
42. Сергиенко, А. Б. Цифровая обработка сигналов: учеб. пособие / А. Б. Сергиенко. — СПб. : БХВ-Петербург, 2011. — 768 с. — ISBN 978-5-9775-0606-9